

**Analyse des Ebolavirus 3'-Leader-Promotors  
– mechanistische Beiträge von VP30, Promotorarchitektur,  
RNA-Sequenzelementen und RNA-Sekundärstrukturen zur  
viralen Transkriptions- und Replikationsinitiation**

Dissertation  
zur Erlangung des Doktorgrades  
der Naturwissenschaften (Dr. rer. nat.)

dem

Fachbereich Pharmazie der  
Philipps-Universität Marburg  
vorgelegt von

**Apothekerin Simone Bach**  
aus **Detmold**

Marburg (Lahn) **2019**

Erstgutachter: **Prof. Dr. Roland K. Hartmann**

Zweitgutachter: **Prof. Dr. Stephan Becker**

Eingereicht am 07.11.2019

Tag der mündlichen Prüfung am 20.12.2019

Hochschulkennziffer: 1180





## ZUSAMMENFASSUNG

In den letzten Jahren sorgte das humanpathogene Ebola-Virus (EBOV) vermehrt für Epidemien vor allem in Zentral- und Westafrika mit Letalitätsraten von bis zu 90 %. Wirksame antivirale Kausaltherapien sind bis dato nicht verfügbar. Aktuell ist die Demokratische Republik Kongo von einer schweren EBOV-Epidemie betroffen.

Das EBOV-Genom ist eine nicht-segmentierte Negativstrang (NNS)-RNA. An den terminalen 3'- und 5'-Genomenden befinden sich regulatorische Elemente der EBOV-Replikation und -Transkription (3'-Leader, 5'-Trailer). Die EBOV-Replikation wird durch einen Proteinkomplex aus drei viralen Proteinen gesteuert, der RNA-abhängigen RNA-Polymerase L, ihrem Kofaktor, dem viralen Protein 35 (VP35), und dem Nukleoprotein (NP). Die virale Transkription benötigt darüber hinaus das virale Protein 30 (VP30) als Transkriptionsaktivator. Vermutlich nutzt die Virus-Polymerase eine einzige Eintrittsstelle am 3'-Genomende, von der aus sowohl Replikation als auch Transkription initiiert werden. Im Falle anderer NNS-Viren führt dies zur Synthese eines kurzen Abbruchtranskripts antisense zum 3'-Leader vor der Initiation der mRNA-Synthese am ersten Genstartsignal.

Der genomische Replikationspromoter ist zweigeteilt. Promoterelemente 1 und 2 (PE1, PE2) sind durch die Transkriptionsstartsequenz (TSS) und weitere Platzhalternukleotide (Spacer) getrennt. Entsprechend der „Rule of 6“ darf der Spacer lediglich um ein Vielfaches der Zahl 6 erweitert oder verkürzt werden, um eine effiziente Replikation zu gewährleisten. PE2 weist darüber hinaus 8 konsekutive 3'-UN<sub>5</sub>-Hexamere auf, von denen 3 obligat für eine residuale Replikationsaktivität sind. Auch der TSS-Spacer-Bereich enthält UN<sub>5</sub>-Hexamere. Diese sind jedoch an Nt -75 durch ein Guanosin anstelle eines Uridins vom PE2-UN<sub>5</sub>-Raster getrennt. Der Spacer kann sowohl auf genomischer als auch auf der komplementären mRNA-Ebene putative Haarnadelstrukturen unter Einbeziehung der TSS ausbilden (NP-Hairpin). Diesem NP-Hairpin wird eine Rolle in der VP30-abhängigen Transkriptionsregulation zugeschrieben.

Ziel der vorliegenden Studie war es, ein besseres Verständnis der Polymerisationsinitiation am 3'-Leader-Promotor zu erlangen. Im Fokus standen neben der Untersuchung der Replikations- und Transkriptionspromotorarchitektur die Analyse regulatorischer RNA-Sequenzen, die Rolle der TSS-Hairpin-Struktur sowie die Regulation der Transkription durch VP30.

Mithilfe einer gezielten Verwendung Replikations-kompetenter und Replikations-defizienter Minigenomvarianten in Kombination mit Reporterassays und qRT-PCR konnten wir demonstrieren, dass PE1 ebenfalls den Transkriptionspromotor kodiert. Darüber hinaus zeigen unsere Daten, dass die "Rule of 6" in der Spacer-Region nicht

nur für die Replikations- sondern auch für die Transkriptionsinitiation relevant ist. Eine Verlängerung des Spacer-Bereichs ist, abhängig vom Sequenz- und Strukturkontext, um mindestens 66 Nt ohne vollständigen Funktionsverlust möglich.

Die Funktion der UN<sub>5</sub>-Hexamere in Spacer und PE2 ist nach wie vor spekulativ. Möglicherweise positionieren die Hexamere das Nukleoprotein während der RNA-Synthese in spezifischer Anordnung auf dem RNA-Templat, um eine produktive Promotorerkennung durch die virale Polymerase zu ermöglichen, oder sie koordinieren die NP-Dissoziation/Reassoziierung am RNA-Templatstrang, während die RNA durch das aktive Zentrum der Polymerase gefädelt wird. Unsere Daten deuten darauf hin, dass ein kontinuierliches UN<sub>5</sub>-Hexamerraster zwischen PE1 und PE2 aufgrund einer G-75 zu U-Punktmutation die allgemeine Effizienz von Transkription und Replikation steigert, jedoch auf Kosten der selektiven Feinregulation beider Prozesse durch VP30 (Transkriptionsaktivierung und Replikationsrepression).

Im Fall einer Spacer-Deletion von 12 Nt, die gleichzeitig die Ausbildung des NP-Hairpins verhindert, ist die Replikations- und Transkriptionsaktivität nach wie vor > 50 % und die Regulation durch VP30 bleibt grundsätzlich erhalten. Dies verdeutlicht, dass Sekundärstrukturen am Transkriptionsstart weder essenziell für die VP30-abhängige Transkription noch für die Replikation sind. Allerdings konnten wir zeigen, dass Hairpinstabilisierungen tendenziell die VP30-Abhängigkeit der Transkription relaxieren, während Stabilisierungen diese steigern. Eine schrittweise Stabilisierung der TSS-Hairpinstruktur durch strategische Insertion von G:C-Basenpaaren und die damit verbundene Ausweitung lokaler Doppelstrangbereiche führte, abhängig vom Ausmaß der Stabilisierung auf Antigenom/mRNA-Ebene, zu starken Aktivitätsverlusten bis hin zur Elimination der Transkriptions- als auch Replikationsaktivität. Damit konnten wir zum ersten Mal indirekt die Ausbildung von RNA-Strukturen während der viralen RNA-Synthese nachweisen. Allgemein führten jedwede Änderungen der Sequenz, Struktur, Stabilität oder Länge des nativen NP-Hairpins zu einer partiell verringerten VP30-Abhängigkeit der Transkription. Somit erscheint der NP-Hairpin optimiert für eine engmaschige Regulation durch VP30.

Hochdurchsatz-RNA-Sequenzierung und Northern Blot-Analyse bestätigten darüber hinaus die Synthese abortiver *leader* RNAs. Ihre Initiation erfolgte wie im Fall der Antigenomsynthese ausschließlich an Nt -2. Im Gegensatz dazu sorgte der Erhalt des ersten Genomnukleotids für eine Vervierfachung der Transkription. Sowohl die *leader* RNA-Initiation an Nt -2 als auch die Termination in einem Längenbereich von ~ 60 - 80 Nt erfolgten unabhängig von RNA-Sequenz/-Struktur oder von VP30. In Anwesenheit von VP30 sind die *leader* RNA-Spiegel allgemein reduziert. 24 h nach EBOV-Infektion bzw. 48 h nach Minigenom-Transfektion sind die *leader* RNA-Spiegel 9-fach bzw. 68-fach geringer als die der ersten mRNA. Zusammenfassend deuten

unsere Daten darauf hin, dass *leader*RNAs Abbruchprodukte der Antigenomsynthese und möglicherweise keine obligaten Vorlauf-Produkte der viralen Transkription sind.

## ABSTRACT

In recent years the human pathogen Ebola virus (EBOV) has been the causative agent of large epidemics in Central and Western Africa with fatality rates of up to 90 %. Up to date efficacious antiviral treatments are still missing. The Democratic Republic of the Congo is currently threatened by another severe EBOV epidemic.

The EBOV genome is a non-segmented negative sense (NNS) RNA. Regulatory elements for viral replication and transcription are located at its 3'- and 5'-terminal regions (3'-leader, 5'-trailer). EBOV replication is conducted by a protein complex comprising three viral proteins, the RNA-dependent RNA polymerase L, its cofactor viral protein 35 (VP35) and the nucleoprotein NP. Viral transcription requires an additional transcription activator, viral protein 30 (VP30). It is thought that the viral polymerase uses a single entry site at the genome's very 3' end to initiate transcription and replication. In case of some other NNS viruses evidence was provided that this mechanism entails the synthesis of an abortive transcript antisense to the 3'-*leader* prior to initiation of mRNA synthesis at the first transcription start sequence (TSS) which is located at some distance to the genome 3' end.

The genomic replication promoter is bipartite. Promoter elements 1 and 2 (PE1, PE2) are separated by the first TSS and additional nucleotides (spacer). To assure efficient replication, the spacer has to follow the „rule of 6“. Thus, it can only be extended or shortened by a multiple of 6 nucleotides. Furthermore, PE2 comprises 8 consecutive 3'-UN<sub>5</sub> hexamers. Three of them are essential for residual replication activity. Furthermore, UN<sub>5</sub> hexamer phasing was identified in this work to extend to the TSS-spacer region and the 5'-terminal part of PE1. However, here hexamer continuity in the TSS-spacer region is interrupted by a guanosine residue (instead of a uridine) at nt -75. TSS and spacer are predicted to form RNA secondary structures both on the genomic as well as on the mRNA level (NP hairpin). This NP hairpin was previously reported to be involved in VP30-dependent transcription regulation.

The present study aimed to gain insight into initiation of polymerization at the genome 3'-leader promoter. Apart from analyses on the replication and transcription promoter architecture, we focused on the analysis of regulatory RNA sequences, constraints of the TSS hairpin structure as well as on regulation by VP30.

By use of selective replication-competent and replication-deficient minigenome variants combined with reporter gene assays and qRT-PCR, we demonstrated that PE1 also encodes the EBOV transcription promoter. Furthermore, our data shows that the “rule of 6” is not only key to productive replication initiation but to transcription initiation as well. Substantial replication and transcription activity was still detected when the

spacer region was expanded by up to 66 nt, depending on sequence and structure context.

The function of spacer and PE2 UN<sub>5</sub> hexamers remains elusive. Presumably, they are involved in positioning NP on the RNA template in a certain phase that ensures productive promoter recognition by the viral polymerase, or they coordinate NP dissociation/reassociation from the RNA template while threading the RNA through the active site of the polymerase. Making UN<sub>5</sub> hexamer phasing continuous between PE1 and PE2 by introduction of a G<sub>-75</sub> to U point mutation led to an overall increase in viral transcription and replication. Though, this was at the expense of regulatory fine-tuning of both processes by VP30 (transcription activation and replication repression).

A spacer deletion of 12 nt which simultaneously abolished NP hairpin structure formation still allowed replication and transcription activities > 50 % and was basically maintaining the regulation by VP30. This stresses the notion that RNA secondary structures are neither essential to VP30-dependent transcription nor to replication. However, we observed a trend toward relaxed VP30 dependency with destabilized hairpin structures, while stabilizations tightened the VP30 dependency of transcription. An incremental extension of helical double strand regions in the TSS hairpin by strategic insertion of G:C base pairs resulted in severe activity losses or elimination of viral polymerization, depending on the extent of stabilization on antigenomic/mRNA level. Thus, we provide first indirect evidence that TSS RNA secondary structures indeed form during viral RNA synthesis. In general, any deviations from the native NP hairpin with regard to sequence, structure, stability and length led to a relaxed VP30 dependency of transcription. Hence, the NP hairpin appears to be optimized for a balanced combination of tight regulation by VP30 and efficient transcription and replication.

Moreover, RNA sequencing and Northern Blot experiments confirmed the synthesis of abortive *leader*RNAs that are initiated opposite to genome position -2, exactly like replicative RNA (cRNA). In contrast, preservation of the first genome nucleotide increased transcription ~ 4-fold. Both initiation at nt -2 or termination in the range of nt -60 to -80 occurred irrespective of deviations in RNA sequence/structure or the absence of VP30. In the presence of VP30, overall *leader*RNA synthesis was reduced. Steady-state *leader*RNA levels were 9-fold or 68-fold lower than mRNA levels 24 h post EBOV infection or 48 h post minigenome transfection. In conclusion, our data suggest that *leader*RNAs are products of abortive antigenome synthesis rather than obligatory pre-products of viral transcription.

# INHALT

## Zusammenfassung

## Abstract

<b>I.</b>	<b>Abbildungsverzeichnis .....</b>	<b>I</b>
<b>II.</b>	<b>Tabellenverzeichnis .....</b>	<b>II</b>
<b>III.</b>	<b>Abkürzungsverzeichnis.....</b>	<b>IV</b>
<b>1</b>	<b>Einleitung.....</b>	<b>1</b>
1.1	Taxonomie und Epidemiologie des Ebola-Virus .....	1
1.2	Pathogenese der Ebola-Virus-Krankheit .....	5
1.3	Diagnose der Ebola-Virus-Krankheit .....	6
1.4	Behandlungsoptionen und Vakzinierung gegen EBOV.....	7
1.5	Die Struktur des Ebolaviruspartikels.....	9
1.6	Der Filovirus-Vermehrungszyklus.....	11
1.7	Das EBOV-Genom.....	13
1.8	EBOV-Replikation und -Transkription.....	14
1.8.1	<i>Regulatorische Elemente der EBOV-Replikation und -Transkription ....</i>	<i>18</i>
1.8.2	<i>Die Rolle von VP30 in der EBOV-Transkriptionsinitiation .....</i>	<i>24</i>
1.9	EBOV-Minigenomsysteme .....	26
<b>2</b>	<b>Fragestellung.....</b>	<b>29</b>
<b>3</b>	<b>Material und Methoden .....</b>	<b>31</b>
3.1	Material .....	31
3.1.1	<i>Chemikalien .....</i>	<i>31</i>
3.1.2	<i>Verbrauchsmaterialien .....</i>	<i>33</i>
3.1.3	<i>Enzyme und Kits .....</i>	<i>34</i>
3.1.4	<i>Lösungen und Puffer.....</i>	<i>35</i>
3.1.5	<i>Wachstumsmedien.....</i>	<i>36</i>
3.1.6	<i>Zellen .....</i>	<i>37</i>
3.1.7	<i>Viren .....</i>	<i>38</i>
3.1.8	<i>Nukleinsäuren .....</i>	<i>38</i>
3.1.9	<i>Vektoren und rekombinante Plasmide.....</i>	<i>41</i>

3.1.10	Geräte.....	43
3.1.11	Verwendete Software.....	44
3.2	Methoden.....	45
3.2.1	Molekularbiologische Methoden.....	45
3.2.2	Zellbiologische Methoden .....	60
3.2.3	Virologische Methoden .....	63
<b>4</b>	<b>Ergebnisse und Diskussion .....</b>	<b>65</b>
4.1	Die Rolle der „Rule of 6“ für die EBOV-Transkription .....	65
4.2	Einfluss von Sekundärstrukturstabilität und Spacer-Länge auf die EBOV-Transkription und -Replikation .....	107
4.3	Detektion und Charakterisierung von EBOV- <i>leader</i> RNAs .....	138
4.4	Der Einfluss von RNA-Sequenz, -Struktur und Länge der TSS-Spacer-Region auf die VP30-Abhängigkeit der Transkription und die Balance von Transkription und Replikation .....	198
<b>5</b>	<b>Diskussion und Zusammenfassung .....</b>	<b>228</b>
<b>6</b>	<b>Ausblick.....</b>	<b>232</b>
<b>7</b>	<b>Literaturverzeichnis.....</b>	<b>240</b>
<b>8</b>	<b>Anhang .....</b>	<b>254</b>
8.1	Ergänzende Klonierungsstrategien und Primer .....	254
8.2	Plasmidsequenz pA-3E5E-GLuc-GFP .....	255
8.3	Manuskripte und Veröffentlichungen.....	257
8.3.1	Manuskripte .....	257
8.3.2	Veröffentlichungen.....	257
8.4	Konferenzbeiträge.....	258
8.5	Lebenslauf .....	259
8.6	Eidesstattliche Erklärung.....	260
8.7	Danksagung.....	261



## I. ABBILDUNGSVERZEICHNIS

Abb. 1: Taxonomie der <i>Mononegavirales</i> .....	1
Abb. 2: Phylogenie von Filoviren .....	2
Abb. 3: Epidemiologie der Ebolaviren.....	3
Abb. 4: EBOV-Vakzine in klinischer Erprobung .....	9
Abb. 5: Filoviruspartikel .....	10
Abb. 6: Kryo-Elektronenmikroskop-Rekonstruktion des EBOV-Nukleokapsids .....	11
Abb. 7: Der Filovirus-Verbreitungszyklus.....	12
Abb. 8: Das EBOV-Genom.....	14
Abb. 9: Modell der EBOV-Replikation und -Transkription .....	16
Abb. 10: Ebola-Virus-Genom 3'-Replikations- und Transkriptionspromotorbereich.....	21
Abb. 11: Anordnung von PE1 und PE2.....	24
Abb. 12: Monocistronisches EBOV-Minigenom .....	26
Abb. 13: EBOV-Minigenomsysteme .....	28
Abb. 14: Strategien der ortsgerechten <i>in vitro</i> -Mutagenese .....	50
Abb. 15: Minigenom-kodierende Plasmide (monocistronisch und bicistronisch) .....	51
Abb. 16: RNA-Qualitätskontrolle.....	58
Abb. 17: EBOV-spezifische Minigenomsysteme.....	63
Abb. 18: Zusammenfassung regulatorischer Sequenz-/Strukturbereiche des 3'-Leader-Promotors .....	228
Abb. 19: NP-Spacer-Variante mit aufgelöster Sekundärstruktur auf vRNA- und mRNA-Ebene aber ohne Veränderung der Spacer-Länge .....	235
Abb. 20: VP35-Hairpin-Austausch an Cistron 2 in bicistronischen MGs.....	236
Abb. 21: Putative Sekundärstruktur eines Leader-Transkripts (leaderRNA; 73 Nt) .....	237

## II. TABELLENVERZEICHNIS

Tab. 1: EBOV-Genstart- und Genendsignale .....	18
Tab. 2: EBOV-Gengrenzen .....	19
Tab. 3: Chemikalien .....	31
Tab. 4: Verbrauchsmaterialien .....	33
Tab. 5: Enzyme .....	34
Tab. 6: Kits.....	34
Tab. 7: Lösungen .....	35
Tab. 8: Puffer .....	36
Tab. 9: Wachstumsmedien für Bakterienzellen .....	36
Tab. 10: Wachstumsmedien für Säugerzellen .....	37
Tab. 11: Prokaryotische Zellen.....	37
Tab. 12: Eukaryotische Zellen.....	38
Tab. 13: Viren .....	38
Tab. 14: RT-Primer .....	38
Tab. 15: qPCR-Primersets .....	39
Tab. 16: Sequenzierprimer.....	39
Tab. 17: Kolonie-PCR-Primer.....	40
Tab. 18: RNA-Oligonukleotide.....	41
Tab. 19: Rekombinante Plasmide und Vektoren .....	42
Tab. 20: Geräte.....	43
Tab. 21: Verwendete Software .....	44
Tab. 22: Standard PCR.....	45
Tab. 23: Standard PCR-Programm .....	45
Tab. 24: Kolonie-PCR-Ansatz .....	46
Tab. 25: Kolonie-PCR-Programm.....	46
Tab. 26: Pipettierschema der Reversen Transkription.....	47
Tab. 27: Programm der Reversen Transkription.....	47

Tab. 28: qPCR-Pipettierschema.....	48
Tab. 29: qPCR Fast-Cycling-Programm. ....	48
Tab. 30: Ansatz <i>DpnI</i> -Restriktionsverdau .....	52
Tab. 31: Ansatz DNA-Phosphorylierung.....	53
Tab. 32: Ligations-Ansatz.....	53
Tab. 33: Ansatz des DNase I-Verdau in Lösung .....	56
Tab. 34: Ansatz der DNA-Sequenzierungen.....	59
Tab. 35: Pipettierschema Minigenom-Assay.....	64
Tab. 36: Cloning strategies and primers of manuscript 1.....	101
Tab. 37: Cloning strategies and primers manuscript 2.....	133
Tab. 38: Cloning strategies and primers manuscript 3.....	190
Tab. 39: Ergänzende Klonierungsstrategien und Primer.....	254

### III. ABKÜRZUNGSVERZEICHNIS

Sofern möglich wurden SI-Einheiten und international verbindliche Bezeichnungen für Nukleinsäure-Nukleotide verwendet. Abkürzungen chemischer Substanzen sind dem Materialteil zu entnehmen.

Abkürzung	Deutsche Bezeichnung	Englische Bezeichnung
5'-PPP; 5'-ppp	5'-Triphosphat	5'-triphosphate
$\lambda$	griechisch: Lambda; Wellenlänge	lambda; wavelength
%	Prozent	percent
°C	Grad Celsius	degrees Celsius
$\mu\text{g}$	Mikrogramm	microgram
$\mu\text{L}$	Mikroliter	microliter
$\mu\text{m}$	Mikrometer	micrometer
$\mu\text{M}$	Mikromolar [mol/L]	micromolar [mol/l]
Abb.	Abbildung	/
Amp	Ampicillin	ampicillin
bp	Basenpaar	base pair
BSL	Biosicherheitslevel	biosafety level
bzw.	beziehungsweise	/
ca.	circa	/
cRNA	copy-RNA; Antigenom	copy-RNA; antigenomic RNA
$C_T$	Cycle threshold; Zyklus-Schwellenwert	cycle threshold
(d)dH <sub>2</sub> O	(doppelt-)destilliertes Wasser	(double-)distilled water
d.h.	das heißt	/
DNA	Desoxyribonukleinsäure	deoxyribonucleic acid
DNase	Desoxyribonuklease	deoxyribonuclease
dNTP	Desoxyribonukleosidtriphosphat	deoxyribonucleotide triphosphate
ds	doppelsträngig	double-stranded
EBOV	Ebola-Virus; Spezies: <i>Zaire-Ebolavirus</i>	Ebola virus; species: <i>Zaire ebolavirus</i>
<i>E. coli</i>	<i>Escherichia coli</i>	/

<b>Abkürzung</b>	<b>Deutsche Bezeichnung</b>	<b>Englische Bezeichnung</b>
FCS	Fetales Kälberserum	fetal calf serum
Fig.	/	figure
fwd	vorwärts	forward
g	Gramm	Gram
GE	Genende	Gene end
GP/G	Glykoprotein	glycoprotein
GS	Genstart	Gene start
h	Stunde	hour
HP	Hairpin/Haarnadelstruktur	hairpin
IFN	Interferon	interferone
IGR	intergenische Region	intergenic region
kb	Kilobase	kilobase
/eRNA	/leaderRNA, Leader-Transkript	/leaderRNA, leader transcript
LB	/	lysogeny broth (cell culture media)
mA	Milliampere	milliampere
MARV	Marburg-Virus	Marburg virus
MFE	/	minimum free energy
MG	Minigenom	minigenome
min	Minute	minutes
mL	Milliliter	milliliter
mM	Millimolar	millimolar
mRNA	messenger-RNA	messenger-RNA
ng	Nanogramm	nanogram
nm	Nanometer	nanometer
nmol	Nanomolar	nanomolar
NNS	nicht-segmentierter Negativstrang	non-segmented negative sense/strand
NP/N	Nukleoprotein	nucleoprotein
Nt, nt	Nukleotid	nucleotide
P	Phospho-Protein; VP35-Analogon	phospho protein; analogous to VP35

<b>Abkürzung</b>	<b>Deutsche Bezeichnung</b>	<b>Englische Bezeichnung</b>
p.a.	lateinisch: per analysis	/
PAGE	Polyacrylamid-Gelelektrophorese	polyacrylamide gel electrophoresis
PCR	Polymerase-Kettenreaktion	polymerase chain reaction
PE1, PE2	(Replikations-)Promotorelemente 1, 2	promotor elements 1 and 2
pH	lateinisch: potentia hydrogenii	/
pmol	Pikomol	picomol
p.t.	post transfectionem; nach Transfektion	post transfectionem; post transfection
OD	Optische Dichte	/
ORF	Offener Leserahmen	open reading frame
(q)RT-PCR	(quantitative) Reverse-Transkriptase-Polymerase-Kettenreaktion	(quantitative) reverse transcriptase polymerase chain reaction
RABV	Rabies-Virus („Tollwutvirus“)	rabies virus
RC	Replikations/Transkriptions-kompetent	Replication/transcription-competent
RD	Replikations-defizient	replication-deficient
RdRp	RNA-abhängige RNA-Polymerase	RNA-dependent RNA polymerase
rev	rückwärts	reverse
RIG-I	Retinsäure-induzierbares Gen I (intrazellulärer Rezeptor des angeborenen Immunsystems, der virale 5'-PPP-RNA erkennt)	retinoic acid inducible gene I
RNA	Ribonukleinsäure	ribonucleic acid
RNase	Ribonuklease	ribonuclease
rpm	Runden pro Minute	rounds per minute
(H)RSV	Humanes Respiratorisches Synzytial-Virus	human respiratory syncytial virus
RT	Reverse Transkription	reverse transcription
Seq	Sequenzierung	sequencing
SeV	Sendai-Virus	Sendai virus
sGP	<i>soluble GP</i> (lösliches Glykoprotein)	soluble GP
ssGP	<i>small soluble GP</i> (kleines, lösliches Glykoprotein)	small soluble GP
ssRNA	einzelsträngige RNA	single-stranded RNA

<b>Abkürzung</b>	<b>Deutsche Bezeichnung</b>	<b>Englische Bezeichnung</b>
SUDV	Sudan-Virus; Spezies: <i>Sudan-Ebolavirus</i>	Sudan virus; species: <i>Sudan ebolavirus</i>
Tab.	Tabelle	table
TFS	Thermo Fisher Scientific	/
TSS	Transkriptionsstartsequenz	transcription start sequence
TSR	Transkriptionsstartregion	transcription start region
u.a.	unter anderem	/
UTR	untranslatierte Region	untranslated region
UV	Ultraviolett	ultraviolet
Vgl.	Vergleich	/
(i)VLP	(infektiöser) Virus-artiger Partikel	(infectious) virus-like particle
VP24	Virales Protein 24; Nukleokapsidmaturierungsfaktor	viral protein 24
VP30	Virales Protein 30; Transkriptionsaktivator	viral protein 30
VP35	Virales Protein 35; Polymerase-Kofaktor	viral protein 35
VP40	Virales Protein 40; Matrixprotein	viral protein 40
vRNA	Virus-RNA; Genom	viral RNA; genomic RNA
vs.	lateinisch versus: gegen	Latin: versus
VS(I)V	Vesikuläres Stomatitis-(Indiana)-Virus	vesicular stomatitis indiana virus
v/v	Volumen pro Volumen [%]	volume to volume [%]
w/v	Masse pro Volumen [%]	weight to volume [%]
WHO	Weltgesundheitsorganisation	World Health Organization
z.B.	Zum Beispiel	/
ZEBOV	veraltet: Zaire-Ebola-Virus; jetzt EBOV	Zaire Ebola virus; now EBOV

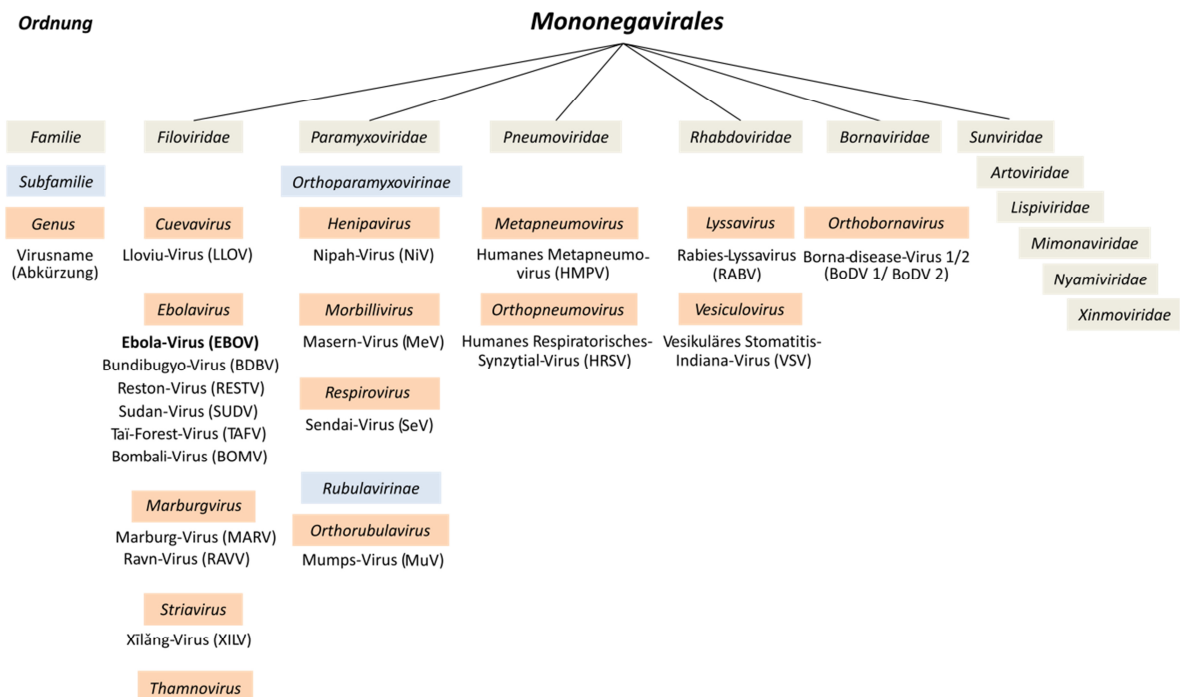




# 1 EINLEITUNG

## 1.1 Taxonomie und Epidemiologie des Ebola-Virus

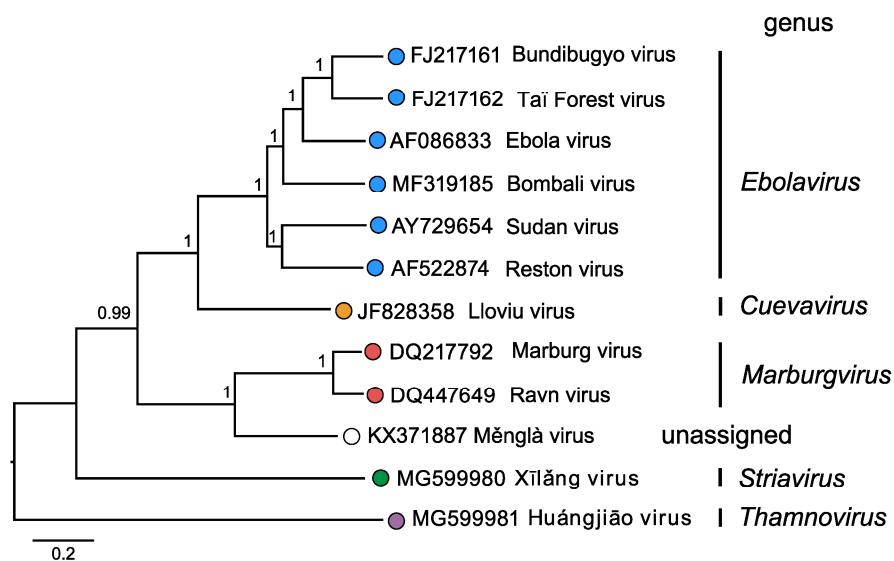
Das Ebola-Virus (EBOV) gehört der Familie der Filoviren (*Filoviridae*) an. Der Name der Virusfamilie lässt sich auf das fadenförmige Aussehen der Viruspartikel (lat. filum = Faden) zurückführen. Das Filovirus-RNA-Genom, welches bei allen Filovirusvertretern eine Länge von 15 bis 19 kb aufweist, ist einzelsträngig, nicht-segmentiert und besitzt eine negative Polarität (Hume and Mühlberger, 2019). Damit gehören die Filoviren zu den NNS-RNA-Viren (NNS = non-segmented negative strand) und lassen sich in die Ordnung der *Mononegavirales* einordnen (Abb. 1; Amarasinghe et al., 2019; Kuhn et al., 2019). Weitere für diese Arbeit erwähnenswerte Vertreter der *Mononegavirales* sind die Familien der *Paramyxoviridae*, *Pneumoviridae*, *Rhabdoviridae* und *Bornaviridae* (Abb. 1).



**Abb. 1: Taxonomie der *Mononegavirales*.** Dargestellt sind Ordnung, Familie, Genus und spezifischer Virusname mit Abkürzung. Virusspezies sind nicht dargestellt und nur für das Genus *Ebolavirus* im Text aufgeführt. Die Familien der *Sunviridae*, *Artoviridae*, *Lispiviridae*, *Mimonaviridae*, *Nyamiviridae* sowie *Xinmoviridae* sind ohne Genera und Virusnamen dargestellt. Im Falle der *Paramyxoviridae*, *Pneumoviridae*, *Rhabdoviridae* und *Bornaviridae* sind jeweils nur die wichtigsten Genera und Virusnamen angegeben. In Anlehnung an (Amarasinghe et al., 2019; Kuhn et al., 2019b).

Zurzeit gehören 5 Genera der Filovirusfamilie an (Abb. 1; Abb. 2): *Ebolavirus*, *Cuevavirus*, *Marburgvirus*, *Striavirus* und *Thamnovirus* (Amarasinghe et al., 2019). Das Genus *Ebolavirus* lässt sich in 6 Spezies unterteilen, die jeweils nach ihrem

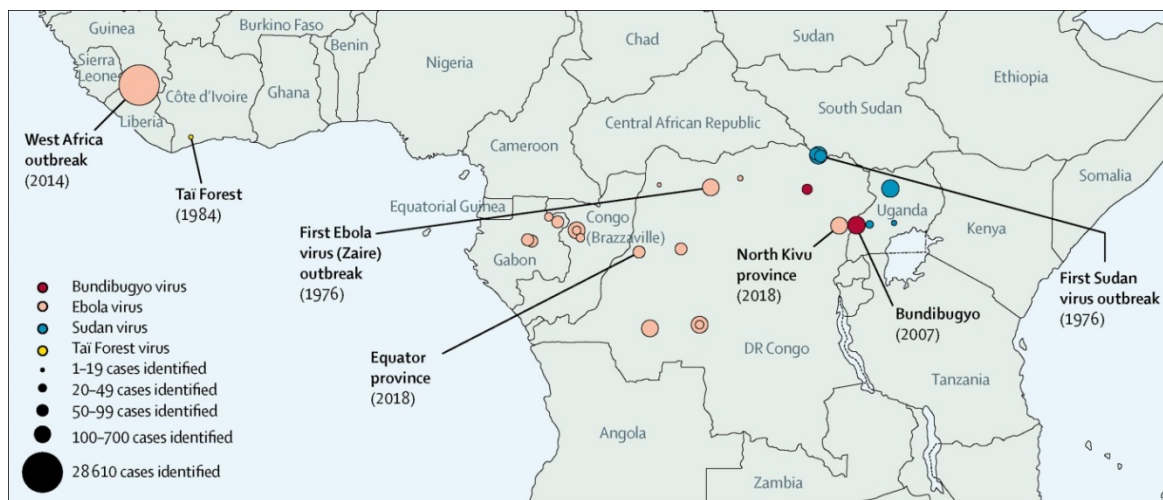
geografischen Fundort benannt sind: *Zaire Ebolavirus* (Virusname: Ebola-Virus, EBOV), *Bundibugyo Ebolavirus* (BDBV), *Reston Ebolavirus* (RESTV), *Sudan Ebolavirus* (SUDV), *Tai Forest Ebolavirus* (TAFV) und *Bombali Ebolavirus* (BOMV). Wobei letzteres bisher nicht mit Infektionen von Menschen in Verbindung gebracht werden konnte, sondern mithilfe von Hochdurchsatzsequenzierung der nächsten Generation (= Next Generation Sequencing, NGS) identifiziert wurde (Goldstein et al., 2018). Gleiches gilt auch für das Lloviu-Virus (LLOV) der Gattung *Cuevavirus*, welches bislang ausschließlich in Langflügelfledermäusen in Spanien und Ungarn nachwiesen werden konnte (Kemenesi et al., 2018; Negredo et al., 2011) sowie für das Mënglà-Virus (MLAV). Dieses wurde in



**Abb. 2: Phylogenie von Filoviren basierend auf dem Vergleich der Sequenz der RNA-abhängigen-RNA-Polymerase (L).** Die Distanzen des phylogenetischen Baums sind auf die evolutionären Abstände nach dem GTR-Modell skaliert [General Time Reversible Model, (Tavaré, 1986)]. Die Bootstrap-Werte resultieren aus 5000 Bootstrap-Replikaten und sind als Dezimalzahl an den Knotenpunkten der Abzweigungen angegeben. Der Bootstrap-Wert ist ein Maß für die statistische Stabilität der Daten und gibt an, wie oft in (in diesem Fall) 5000 Wiederholungen der phylogenetischen Analyse mit zufälligen Teilmengen der Daten dieselbe Abzweigung für dasselbe Datenset ermittelt wurde (1= in allen Replikaten). Vor dem jeweiligen Virusnamen sind die zur Analyse herangezogenen Gendatenbank-Identifikationsnummern angegeben. ICTV-Report, März 2019: [https://talk.ictvonline.org/ictv-reports/ictv\\_online\\_report/negative-sense-rna-viruses/mononegavirales/w/filoviridae](https://talk.ictvonline.org/ictv-reports/ictv_online_report/negative-sense-rna-viruses/mononegavirales/w/filoviridae).

einem Rosettenflughund in China isoliert und wird möglicherweise demnächst als erster Vertreter eines neuen, sechsten, Genus „*Dianlovirus*“ in die Filovirusfamilie aufgenommen (Yang et al., 2019). Von den 6 oben genannten Filovirus-Genera haben *Cuevaviren*, *Ebolaviren* und *Marburgviren* Säugetierwirte. In der Regel sind dies Fledermäuse und Flughunde (Burk et al., 2016). Das für den Menschen ungefährliche REBOV aus dem Genus *Ebolavirus* wurde jedoch auch in Schweinen nachgewiesen (Normile, 2009). Interessanterweise scheinen Vertreter der Stria- und Thamnoviren keine Säugetierwirte, sondern Fischwirte zu nutzen (Shi et al., 2018). Für Menschen (und Affen) sind einige der Filoviren hochpathogen. Dazu gehören das Marburg-Virus

(MARV) sowie die *Ebolavirus*-Vertreter SUDV, BDBV, und EBOV. Gerade das EBOV verursachte in den letzten Jahren vermehrt epidemische Ausbrüche von hämorrhagischem Fieber (= Ebolafieber), vor allem in West- und Äquatorialafrika (Abb. 3). In den meisten Fällen können EBOV-Infektionen beim Menschen initial auf den Kontakt mit oder den Verzehr von Buschfleisch (infizierte Affen oder Flughunde) zurückgeführt werden (Leroy et al., 2004). Sekundäre Mensch-zu-Mensch-Übertragungen erfolgen durch direkten Kontakt mit Blut, Sekreten oder anderen Körperflüssigkeiten infizierter Menschen, entweder bei der Pflege Erkrankter oder der traditionellen Beisetzung von toten Körpern. Ebenso ist bekannt, dass Überlebende der Ebola-Virus-Krankheit das Virus noch Monate oder sogar Jahre nach Symptombefreiheit sexuell übertragen können. Darüber hinaus ist davon auszugehen, dass einige Ebolavirusinfektionen völlig symptomfrei erfolgen und damit unerkannt bleiben [Zusammengefasst in: (Malvy et al., 2019)].



**Abb. 3: Epidemiologie der Ebolaviren** (Malvy et al., 2019).

Die erste humane Infektion mit Filoviren trat im Jahr 1967 auf. Damals erkrankten einige Labormitarbeiter in der Impfstoff- und Serenproduktion in Marburg, Frankfurt und Belgrad an einem schweren und für einige der Erkrankten (ca. 23 %) letalen hämorrhagischen Fieber. Sie waren bei pharmakologischen Experimenten an aus Uganda importierten Primaten der Gattung Grüne Meerkatze (*Cercopithecus aethiops*) mit dem damals noch unbeschriebenen Marburg-Virus (MARV) in Berührung gekommen (Martini, 1971). Nur wenige Monate nach Ausbruch der Infektion konnte das Marburg-Virus erfolgreich im Elektronenmikroskop dargestellt und charakterisiert werden. Es wurde nach der Stadt seiner Erstbeschreibung, Marburg, benannt. Nachfolgend kam es zu weiteren sporadischen, zum Teil aber auch größeren MARV-Ausbrüchen in Angola, der Demokratischen Republik Kongo, Kenia, Südafrika und Uganda mit Letalitätsraten von im Schnitt 80,8 % (Burk et al., 2016).

Ebolaviren wurden erstmalig 1976 während zweier nahezu zeitgleicher Ausbrüche im Sudan (heute: Südsudan) und in der Region Zaire (Heute: Demokratische Republik Kongo) entdeckt (Abb. 3; Johnson et al., 1977). Früh war klar, dass die isolierten Viruspartikel in beiden Fällen morphologisch zwar dem MARV ähnelten, sich immunologisch jedoch von diesem unterschieden und somit eine neue Virusgattung darstellten, die später *Ebolavirus* genannt wurde (Brès, 1978; Commission, 1978; Deng et al., 1978). Das Virus des ersten Ausbruchs im Sudan, mit einer Letalitätsrate von rund 53 %, erhielt den Virusnamen Sudan-Virus (SUDV). Das in Zaire isolierte Virus erhielt seinen Namen Ebola-Virus (EBOV) nach dem im Endemiegebiet verlaufenden Fluss, Ebola. Hier lag die Letalitätsrate bei 88 % (Rougeron et al., 2015). Dem ersten Ebolavirusausbruch folgten etwa 20 weitere Ausbrüche, vor allem im Sudan, in Uganda, in der Demokratischen Republik Kongo und im Staat Gabon (Abb.3). Sie ließen sich hauptsächlich auf Infektionen durch SUDV oder EBOV zurückführen. Es ist durchaus möglich, dass einige weitere kleinere Ausbrüche unidentifiziert blieben. Während die meisten der Ausbrüche auf ländliche Gebiete beschränkt waren, so war beim Ausbruch in Gulu im Jahr 2000 eine semi-urbane Region in Uganda betroffen (Malvy et al., 2019). Der bis heute größte, endemische Ausbruch von Ebolaviren, verursacht durch EBOV, fand 2013 bis 2016 in gleich mehreren Westafrikanischen Staaten statt, darunter Guinea, Sierra Leone und Liberia (Abb.3). Dieses Mal waren sowohl ländliche als auch städtische Gebiete mit einer Fallzahl von > 28 000 und > 11 000 Todesfällen betroffen (Briand et al., 2014; Lo et al., 2017). Es ist jedoch davon auszugehen, dass die Dunkelziffer deutlich höher lag. Die mittlere Letalitätsrate aller klinisch bestätigten Fälle lag bei ca. 63 % (Garske et al., 2017). Im Mai und August 2018 gab es erneut EBOV-Ausbrüche im Kongo (Mbala-Kingebeni et al., 2019). Das erste Ausbruchsgebiet betraf die abgeschiedene Äquatorprovinz. Der Ausbruch wurde erst Anfang Mai publik, nachdem dieser das Stadtgebiet Mbandakas erreichte, einen wichtigen Verkehrsknotenpunkt nahe des Kongo-Flusses (Malvy et al., 2019). Der zweite Ausbruch folgte im August in der Nord Kivu-Provinz (Abb. 3) und weitete sich später auf die benachbarte Ituri-Provinz aus. Während der erste Ausbruch nach insgesamt 54 Fällen, und davon 33 Todesfällen, eingedämmt werden konnte, ist der Ausbruch in den Provinzen Nord Kivu und Ituri bis heute nicht vollständig unter Kontrolle, da die medizinische Versorgung sowie Präventionsmaßnahmen durch gewaltsame Konflikte in der betroffenen Region erschwert werden (Malvy et al., 2019). Der aktuelle „External Situation Report 66“ der WHO vom 05.11.2019 berichtet von bislang insgesamt 3274 Fällen. Darunter sind 3157 bestätigte Fälle mit insgesamt 2185 Todesfällen (World Health Organization, 2019). Damit handelt es sich bereits jetzt um den zweitgrößten EBOV-Ausbruch der Geschichte.

## 1.2 Pathogenese der Ebola-Virus-Krankheit

Die ersten Symptome der Ebola-Virus-Krankheit treten abrupt nach einer Inkubationszeit von 2-21 Tagen auf. Sie sind jedoch unspezifisch (Fieber, Unwohlsein, Schwächegefühl, Müdigkeit und Kopfschmerz; Malvy et al., 2019). Darauf folgen ca. 4-6 Tage nach Erstsymptomatik gastrointestinale Symptome wie Appetitlosigkeit, Bauchschmerzen, Übelkeit, Erbrechen und Durchfall, die je nach Schwere des Verlaufs zu einem Flüssigkeitsverlust von bis zu 5-10 L pro Tag führen können (Malvy et al., 2019). Innerhalb der ersten Woche ist darüber hinaus bei einigen Patienten ein makulopapulöses Exanthem möglich (Leligdowicz et al., 2016). In einigen Fällen wird von Schluckauf berichtet. Weiterhin wird von neurologischen Symptomen wie Delirium, Verwirrung und Hirnhautentzündungen berichtet (West and von Saint André-von Arnim, 2014). Vor allem bedingt durch einen hohen Flüssigkeitsverlust während der gastrointestinalen Symptome und durch eine erhöhte Endothelpermeabilität aufgrund einer überschießenden Immunreaktion sind Hypokaliämie, eine Metabolische Azidose, akute Nierenschädigungen und Nierenversagen sowie ein hypovolämischer Schock möglich (Leligdowicz et al., 2016). Schwere hämorrhagische Komplikationen sind selten, wohingegen in etwa 30 % aller Fälle milde Blutungen an Schleimhäuten, insbesondere des Gastrointestinaltrakts, auftreten können (Leligdowicz et al., 2016). Ist ein virales hämorrhagisches Fieber manifestiert, treten schwerwiegende Blutgerinnungsstörungen auf, die zu Organnekrosen der Lymphknoten, Milz und Leber (frühes Stadium) sowie der Nieren und der Lunge (spätes Stadium) führen können. Bei besonders fulminantem Verlauf kommt es 6-16 Tage nach Erstsymptomatik zum Tod durch Multiorganversagen und hypovolämischem Schock (Basler, 2017; Leroy et al., 2011).

Die ersten von Ebolaviren infizierten Zellen sind antigenpräsentierende Zellen (Makrophagen, Monozyten und dendritische Zellen) des Lymphgewebes, u.a. der Lymphknoten. Im Laufe der Infektion werden jedoch weitere Körperzellen befallen, darunter Leberzellen und Endothelzellen der Blutgefäße (McElroy et al., 2018). Interessanterweise unterdrückt die Ebolavirus-Infektion die angeborene Immunantwort in vielen Zelltypen, beispielsweise durch Unterbinden der Interferon (IFN)-Freisetzung durch die Virusproteine (VP) 35 (RIG-I-vermittelt) und 24 (VP24) sowie das Oberflächenglykoprotein GP (McElroy et al., 2018). Makrophagen bilden bemerkenswerter Weise die Ausnahme. Die Aktivierung von Makrophagen in der frühen Infektionsphase führt zur massiven Freisetzung von Zytokinen [IFN, Interleukine (IL)-2, -6, -8 und -10, Tumornekrosefaktor alpha (TNF- $\alpha$ )] (Basler, 2017). Darüber hinaus kann es zu einem Verlust eines Teils der Monozyten kommen, jedoch ist noch unklar, ob dies über Apoptose geschieht (Malvy et al., 2019). In den späteren Phasen der Infektion wird

auch das Endothel infiziert. Dies führt ebenfalls zur Freisetzung inflammatorischer Mediatoren. Die exzessive Freisetzung von Zytokinen (Zytokinsturm) und die unkontrollierte Virusvermehrung begünstigen Gefäßläsionen, bzw. erhöhen die Endothelpermeabilität (Durchlässigkeit). Auch die intravaskuläre Blutgerinnung wird gestört (Basler, 2017; Leligidowicz et al., 2016; Malvy et al., 2019). In fatalen Fällen ist die adaptive Immunantwort ebenfalls stark beeinträchtigt, vermutlich aufgrund VP35- und VP24-vermittelter fehlerhafter Maturierung dendritischer Zellen (Zusammengefasst in McElroy et al., 2018). Dendritische Zellen werden im Gegensatz zu Monozyten und Makrophagen während der Infektion nicht aktiviert. Infolgedessen bleibt die T-Zell-Aktivierung aus. Zusätzlich gehen T-Zellen und Natürliche Killerzellen vermehrt in die Apoptose, ohne selbst direkt infiziert zu werden (Basler, 2017; Bosio et al., 2003; Reed et al., 2004). Auch wenn die zugrundeliegenden Mechanismen und Signalwege noch nicht vollständig aufgeklärt sind, ist davon auszugehen, dass der exzessive pro-inflammatorische Zytokinsturm und das Fehlen der adaptiven Immunantwort maßgeblich für die Ebolavirus-Pathogenese sind (Basler, 2017). Das Ausbleiben der Antigenpräsentation dendritischer Zellen mag der Grund sein, warum Patienten mit fatalem hämorrhagischem Fieber eine persistierende hohe Viruslast in Abwesenheit von virusspezifischen Antikörpern aufweisen, während in Überlebenden im Gegensatz dazu eine potente adaptive Immunantwort nachweisbar ist (Ksiazek et al., 1999; McElroy et al., 2015).

### **1.3 Diagnose der Ebola-Virus-Krankheit**

Bereits 3-6 Tage nach Auftreten der ersten Symptome sind Ebolaviren mithilfe der Reverse-Transkriptase-Polymerase-Kettenreaktion (RT-PCR) im Patientenblut nachweisbar (Towner et al., 2004). Post mortem wird eine RT-PCR des Oralabstrichs durchgeführt. Die RT-PCR-Diagnostik ersetzte während des bislang größten Ebolavirusausbruchs 2013-2016 den wesentlich aufwändigeren und zeitintensiveren Virusnachweis per Anzucht in Zellkulturen (Zusammengefasst in: (Malvy et al., 2019)). Während der akuten Infektion oder auch in der Konvaleszenz-Phase kann Virus-RNA durch RT-PCR in sämtlichen Körperflüssigkeiten nachgewiesen werden, u.a. im Speichel, in der Tränenflüssigkeit, im Schweiß, in der Muttermilch, im Urin sowie im Vaginalsekret oder in der Samenflüssigkeit (Malvy et al., 2019). Erfolgt der Krankheitsverlauf asymptomatisch, so ist die serologische Untersuchung von Immunglobulinspiegeln (IgG und IgM) ca. drei Wochen nach der vermeintlichen Infektion die diagnostische Methode der Wahl (Glynn et al., 2017). Unabhängig von der Schwere des Krankheitsverlaufs, neigen Ebolaviren dazu, im Körper Infizierter zu persistieren, v.a. im Auge, im zentralen Nervensystem oder im Hoden, das heißt an Orten, die für das

Immunsystem schwer zugänglich sind, bzw. die per Definition in der Lage sind, fremde Antigene zu tolerieren (McElroy et al., 2018). Ebolaviren sind in der männlichen Samenflüssigkeit somit noch Monate bis Jahre nach einer akuten Infektion sexuell übertragbar (Deen et al., 2017; Diallo et al., 2016).

## **1.4 Behandlungsoptionen und Vakzinierung gegen EBOV**

Bis dato gibt es keine zugelassene medizinische Kausaltherapie gegen EBOV. Die meisten antiviralen Therapien, an denen bisher geforscht wurde, zielten darauf ab, die virale Replikation, das heißt die Herstellung von Genomkopien, zu senken. Auf diese Weise soll der inflammatorische Zytokinsturm verhindert und damit das Immunsystem in die Lage versetzt werden, das Virus erfolgreich zu bekämpfen. Einige dieser Therapieansätze, die zuvor in Tierversuchen eine hohe Wirksamkeit zeigten, konnten während des EBOV-Ausbruchs 2013-2016 getestet werden (Malvy et al., 2019). Darunter war der small-molecule-Inhibitor Favipavir (Toyama Chemical, Japan). Hierbei handelt es sich um ein Nukleosidanalogen, das die virale Polymerase inhibiert. Ursprünglich wurde es für die Therapie von Influenzavirusinfektionen entwickelt. Auch ein Nukleinsäure-basiertes Therapeutikum (TKM-130803, Arbutus Biopharma, Kanada) gegen die viralen mRNAs der Polymerase L und des Polymeraseko-faktors VP35 (= virales Protein 35) wurde während des Ausbruchs in Westafrika getestet. Eine weitere Therapieoption stellte das Plasma konvaleszenter Spender mit jedoch unbekannten Mengen neutralisierender Antikörper dar. Allerdings zeigte keines der investigativen Therapeutika eine Wirkung gegen das Virus, bzw. bot einen signifikanten Überlebensvorteil [Zusammengefasst in (Malvy et al., 2019)]. Vielversprechend war der Einsatz des Antikörper-Cocktails ZMapp (MappBio, USA), bestehend aus drei humanisierten monoklonalen Antikörpern gegen Epitope des viralen Glykoproteins. Leider konnte die randomisierte, kontrollierte Studie während des Ausbruchs nicht abgeschlossen werden und war somit nicht aussagekräftig. Dies liegt darin begründet, dass zum Ende des Ausbruchs nicht mehr die statistisch erforderliche Zielgröße an Patienten erreicht werden konnte (Davey et al., 2016). Derzeit werden während des aktuellen Ausbruchs in der North Kivu Provinz in der Demokratischen Republik Kongo (bestätigt im August 2018), erneut ZMapp sowie das Nukleosidanalogen Remdesevir (GS-5734), der monoklonale Antikörper mAb114 gegen das virale Glykoprotein und ein koformulierter Cocktail aus drei humanisierten monoklonalen Antikörpern (REGN3470-3471-3479) gegen drei nicht-überlappende Epitope des EBOV-Glykoproteins in einer randomisierten, kontrollierten Studie (NCT03719586) getestet (Gaudinski et al., 2019; Nakkazi, 2018; Sivapalasingam et al., 2018).

Während antivirale Therapien oder eine passive Immunisierung bislang ohne eindeutigen Nutzen blieben, gab es bereits vor dem endemischen EBOV-Ausbruch in Westafrika in den Jahren 2013-2016 vielversprechende Vakzinierungsansätze, die dann während des Ausbruchs, in ihrer Ausentwicklung stark beschleunigt wurden (Malvy et al., 2019), auch aufgrund des hohen medialen Interesses. Der sicherlich am besten präklinisch in Tiermodellen (Rhesusaffe, Maus, Meerschweinchen, Schwein) untersuchte Impfstoff war zu diesem Zeitpunkt das rekombinante Vakzin rVSV-ZEBOV (Abb. 4), bestehend aus Replikations-kompetenten (= replikationsfähigen) Vesikulären Stomatitis-Indiana-Viren (VSV) und in ihre Hülle integrierten EBOV-Glykoproteinen. Aus diesem Grund wurde die rVSV-ZEBOV-Vakzine in Studien der Klinischen Phasen 1-3 aufgenommen und hinsichtlich ihrer Sicherheit und Wirksamkeit im Menschen getestet (Suder et al., 2018). Die Ende 2014 in den USA und der Schweiz begonnenen ersten Phase 1- und 2-Studien zeigten ein hohes immunologisches Potenzial nach Gabe einer Einzeldosis und eine insgesamt gute Verträglichkeit der Vakzine. Allerdings deutete die Studie in der Schweiz an, dass möglicherweise die Ausbildung einer Arthritis mit Gabe der Vakzine korrelierte und zwar auch bei verringerter Dosis. Diese Vakzinnebenwirkung konnte jedoch nicht in den Parallel- und Folgestudien während der Epidemie in Afrika bestätigt werden (Malvy et al., 2019; Suder et al., 2018). Drei Phase 3-Studien wurden durchgeführt, davon zwei 2015 in Guinea und Sierra Leone in Form von Ringvakzinierungen von Erkrankten, Kontaktpersonen sowie Kontaktpersonen der Kontaktpersonen. Die dritte Phase 3-Studie erfolgte als multizentrische Studie in den USA, Spanien und Kanada. In allen Fällen erwies sich die Vakzine als schnell und effizient im Menschen und verursachte lediglich milde unerwünschte Nebenwirkungen oder Komplikationen (grippeähnliche Symptome, in wenigen oder Einzelfällen Fieber oder allergische Reaktionen). Aus diesem Grund wurde die Vakzine 2018 ebenfalls während des Ausbruchs in der Demokratischen Republik Kongo in Ringvakzinierungen eingesetzt. Bis Anfang November 2019 wurden systematisch über 247.000 Menschen geimpft, insbesondere medizinisches Personal, Einsatzkräfte und Kontaktpersonen von Erkrankten (Mitteilung der WHO, 06.11.2019). Die baldige Zulassung des rVSV-ZEBOV-Impfstoffs durch die Europäische Arzneimittel-Agentur (EMA) ist wahrscheinlich. Eine Zulassungsempfehlung der EMA liegt seit 18.10.2019 vor. Neben dieser VSV-Vakzine gibt es weitere vielversprechende Ansätze, die auf adenoviralen Vektoren basieren (Abb. 4) und ebenfalls während des Ausbruchs in Westafrika in Studien der Phasen 2-3 getestet wurden (Malvy et al., 2019). Ab Mitte November 2019 soll laut Mitteilung der WHO (Stand: 06.11.2019) die adenovirale EBOV-Vakzine von Johnson & Johnson (Ad26.ZEBOV, siehe Abb. 4) im aktuellen Ausbruch in der Nord Kivu Provinz als zweite experimentelle Vakzine neben rVSV-ZEBOV eingesetzt werden. Weitere



Vakzinkandidaten basierend auf dem 2013-2016er EBOV-Stamm wurden zwischenzeitlich in China und Russland in klinische Studien aufgenommen (Abb. 4). Der chinesische, adenovirusbasierte Ansatz ist in seiner lyophilisierten (gefriergetrockneten) Form mehr als zwei Wochen auch bei Lagerungstemperaturen von 37 °C stabil. Diese Vakzine wurde bereits in einer Phase 1 und 2 Studien in gesunden Erwachsenen in China und Sierra Leone positiv evaluiert, bedarf aber noch zukünftiger weiterer klinischer Testung (Zhu et al., 2017).

	Leading company or institution and country of origin	Vector	Administration	Ebola component and glycoprotein*	Comments
Recombinant VSV-ZEBOV	Merck (USA)	VSV	Single dose	Ebola virus, Kikwit strain (1995)	Replication competent vaccine
ChAd3-EBO-Z with or without MVA-BN-Filo	GlaxoSmithKline (UK) and, for MVA-BN-Filo, Bavarian Nordic (Denmark)	Chimpanzee adenoviral serotype 3 or MVA	Single dose or heterologous prime-boost regimen	Ebola virus, Mayinga strain (1976)	..
Ad26.ZEBOV with MVA-BN-Filo	Johnson & Johnson (USA), and MVA-BN-Filo from Bavarian Nordic (Denmark)	Human adenoviral serotype 26 or MVA	Heterologous prime-boost regimen	Ebola virus, Mayinga strain (1976)	MVA-BN-Filo encodes Ebola virus, Sudan virus, and Marburg virus glycoproteins, and Tai Forest virus nucleoprotein
Ad5-ZEBOV	Academy of Military Medical Sciences and CanSino Biologics (China)	Human adenoviral serotype 5	Single dose or homologous prime-boost regimen	Ebola virus, Makona strain (2014)	Freeze-dried powder, stable for more than 2 weeks even if kept at a temperature of 37°C; licensed in China
GamEvac-Combi	Gamalei Scientific Research Institute of Epidemiology and Microbiology (Russia)	VSV and Ad5-vectored vaccine	Heterologous prime-boost regimen	Ebola virus, Makona strain (2014)	Licensed in Russia

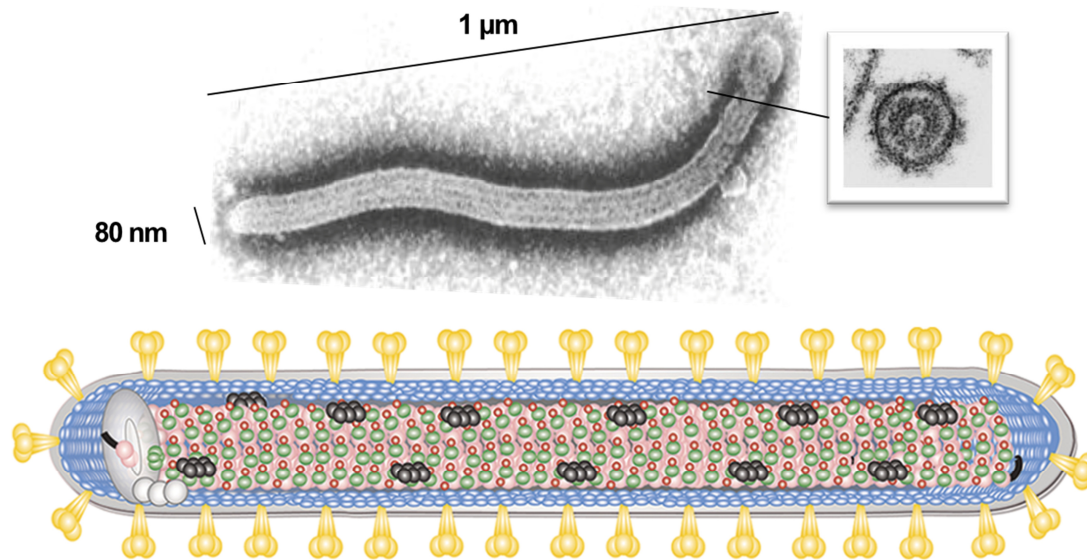
VSV=vesicular stomatitis Indiana virus. MVA=modified vaccinia Ankara virus. Ad5=human adenoviral serotype 5. \*The year the strain (from which the glycoprotein was derived) was isolated is given in brackets.

**Abb. 4: EBOV-Vakzine in klinischer Erprobung während des Ausbruchs 2013-2016** (Malvy et al., 2019).

## 1.5 Die Struktur des Ebolaviruspartikels

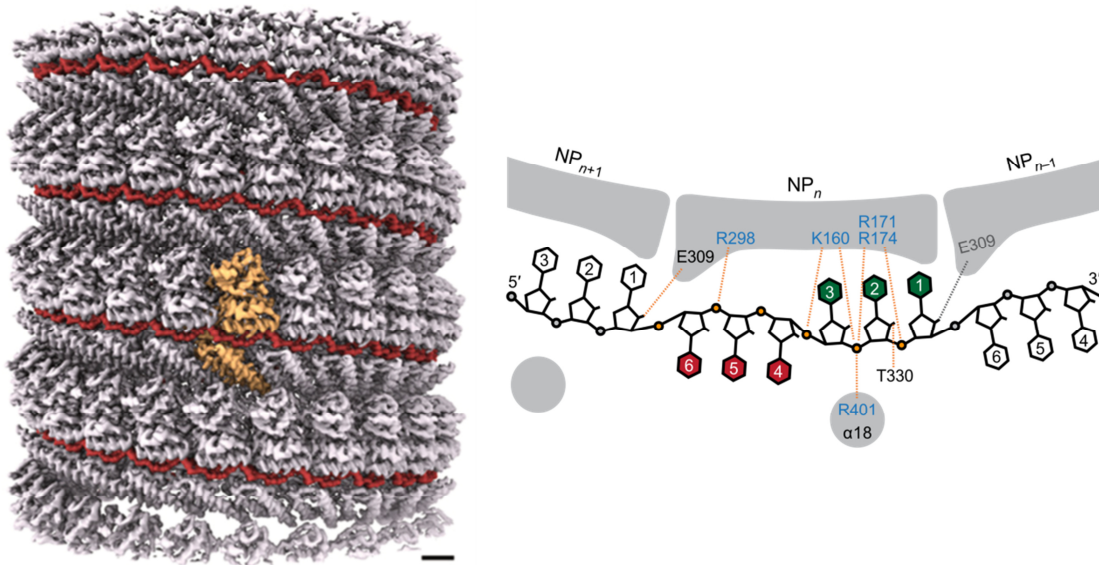
Filovirenpartikel sind fadenförmig mit einer Länge von etwa 1 bis 1.4 µm und einem Durchmesser von etwa 80 nm (Abb. 5). In der Regel liegen sie längsgestreckt vor, können aber auch wie ein Ring oder wie die Zahl 6 geformt sein (Davey et al., 2017; Geisbert and Jahrling, 1995; Peters et al., 1971). Die äußere Virushülle ist aus einer Membran des Wirtsorganismus' geformt, die beim Ausknospen (budding) des Viruspartikels aus den Wirtszellen durch Zellmembranabschnürungen entsteht (Davey et al., 2017). Das einzige Oberflächenprotein des Filovirus, das Glykoprotein GP, ist in ihr als Trimer verankert. Die GP-Trimere werden aus zwei kovalent gebundenen Untereinheiten GP1 und GP2 (GP<sub>1,2</sub>) gebildet. Diese entstehen posttranslational in der infizierten Zelle, in dem das GP-Vorläuferprotein proteolytisch gespalten wird. Die GP1-Untereinheit dient der Zellrezeptorbindung, d.h. dem Anheften an die Wirtszelle, während die membranständige GP2-Untereinheit die Membranfusion katalysiert (Davey et al., 2017; Lee and Saphire, 2009). Von innen ist die äußere Virushülle dicht mit dem viralen Matrixprotein (virales Protein 40, VP40) ausgekleidet. VP40 ist somit der wichtigste Formgeber des Partikels und zudem maßgeblich an der Partikelassemblierung (*assembly*), d.h. des Partikelzusammenbaus, und am Ausknospen des Viruspartikels beteiligt. So konnte gezeigt werden, dass VP40 allein ausreichend ist, um virusartige Partikel (virus-like particles, VLPs) zu generieren (Hume and Mühlberger, 2019). Den innersten Kern des Partikels bildet das Nukleokapsid.

Dieses besteht aus dem viralen Genom, welches in einer linksgedrehten Helix (Abb. 5) durch die viralen Proteine NP (Nukleoprotein), 24 (VP24; Nukleokapsid-maturierungsfaktor) und 35 (VP35) stark kondensiert verpackt (enkapsidiert) wird (Sugita et al., 2018; Wan et al., 2017). Darüber hinaus sind das virale Protein 30 (VP30, Transkriptionsfaktor) sowie der virale RNA-abhängige RNA-Polymerase-Komplex (RdRp), bestehend aus der katalytischen Proteinuntereinheit L und ihrem tetrameren



**Abb. 5: Filoviruspartikel. Oben:** Elektronenmikroskopaufnahme eines Ebola-Virus (EBOV)-Partikels in 8000-facher Vergrößerung (Dr. Larissa Kolesnikova, Institut für Virologie, Philipps-Universität Marburg). **Unten:** Schematische Darstellung des Inneren eines EBOV-Partikels. In der äußeren Virushülle (hellgrau) ist das Oberflächenprotein des Filovirus, das Glykoprotein GP (gelb), verankert. Von innen ist die Virushülle mit dem viralen Matrixprotein VP40 (blau) ausgekleidet. Der innerste Kern des Partikels (Nukleokapsid) besteht aus dem helikalen viralen Genom (dicke schwarze Linie), welches durch das virale Nukleoprotein NP (rosa) sowie VP24 (rot) und VP35 (grün) enkapsidiert wird. Assoziiert sind des Weiteren das virale Protein VP30 (schwarz) sowie der RNA-abhängige RNA-Polymerasekomplex (RdRp), bestehend aus der katalytischen Untereinheit L (ovaler, hellgrauer Ring mit hypothetischem Appendix aus drei globulären Domänen analog zur VSV-L-Struktur (Liang et al., 2015) und ihrem tetrameren Kofaktor VP35 (Hume and Mühlberger, 2019).

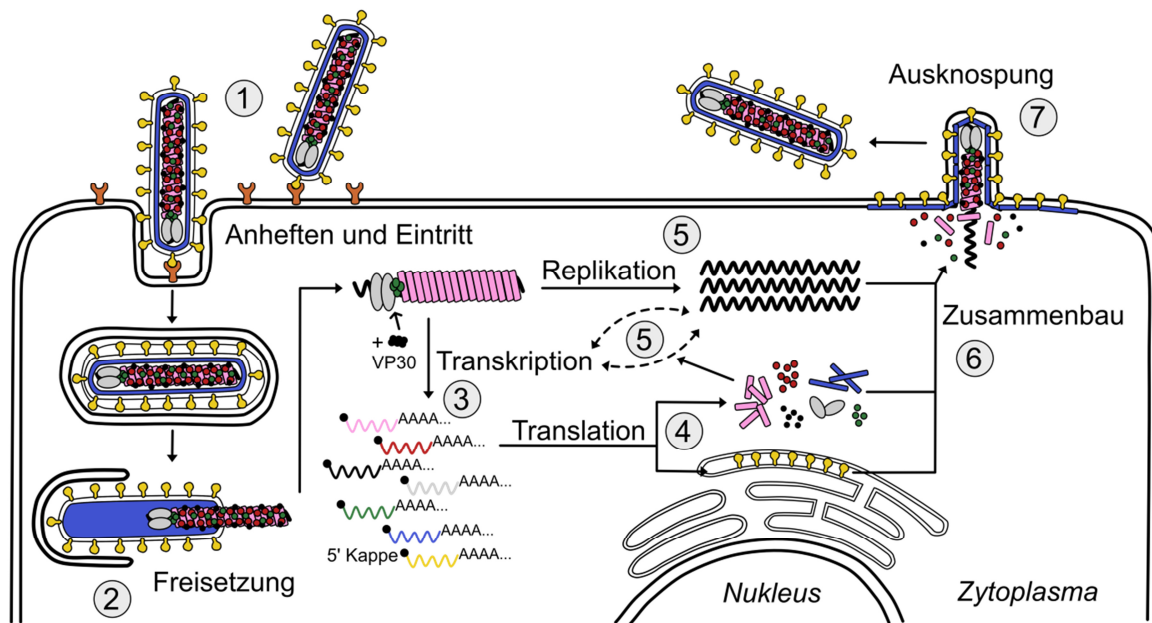
Kofaktor VP35, assoziiert. Pro Nukleokapsidwindung sind im Falle des Ebola-Virus etwa 13-14 NP-Untereinheiten auf dem RNA-Strang assembliert (Wan et al., 2017). Jedes NP-Molekül bindet dabei an 6 RNA-Nukleotide (Abb. 6). Von diesen 6 Nukleotiden sind immer jeweils 3 Nukleotide nach innen zum NP-Kern und 3 Nukleotide davon abgewandt nach außen gerichtet. Die Konfiguration dieser Ausrichtung erfolgt über Wasserstoffbrückenbindungen und elektrostatische Interaktionen zwischen dem NP-Kern und den negativ geladenen RNA-Phosphatgruppen sowie über Wasserstoffbrückenbindungen zwischen RNA-Rückgrat und den polaren Gruppen der äußeren NP- $\alpha$ -Helices (Sugita et al., 2018).



**Abb. 6: Kryo-Elektronenmikroskop-Rekonstruktion des EBOV-Nukleokapsids.** Links: Nukleokapsidkomplex aus viraler RNA und NP-Fragmenten (NP1-450). Der Nukleokapsidkomplex formt eine linksgedrehte helikale Röhre mit einer charakteristischen Reißverschlussartigen Anordnung paralleler C-terminaler  $\alpha$ -Helices. Der RNA-Strang und ein einzelnes NP-Molekül sind farblich in dunkelrot, bzw. in orange hervorgehoben. Rechts: Schema der von NP enkapsidierten RNA. Zwei RNA-Basentriplets mit entgegengesetzter Orientierung werden von einer NP-Untereinheit koordiniert. Die RNA-Basen sind sequentiell in 3'-5'-Orientierung nummeriert und repräsentieren das Negativstranggenom. Nach innen zum NP-Kern ausgerichtete Basen sind in dunkelgrün abgebildet, Basen, die entgegengesetzt, bzw. nach außen ausgerichtet sind, in rot. Die Orientierung der Nukleotide wird durch spezifische elektrostatische Interaktionen (gestrichelte orangene Linien) mit dem NP-Kern und den C-terminalen  $\alpha$ -Helices (hier angezeigt:  $\alpha$ -Helix 18) vermittelt (Sugita et al., 2018).

## 1.6 Der Filovirus-Vermehrungszyklus

Der Filovirus-Lebenszyklus (Abb. 7) lässt sich im Wesentlichen in 7 Phasen unterteilen: (1) Anheften und Eintritt des Viruspartikels an und in die Wirtszelle; (2) Freisetzung des Nukleokapsids in das Zytoplasma; (3) Primärtranskription des Virusgenoms in sieben mRNAs; (4) Translation der mRNAs in virale Proteine; (5) Replikation des Genoms durch neu gebildete Virusproteine sowie sekundäre Transkription basierend auf den neu gebildeten Virusgenomen; Zusammenbau (6) sowie Ausknospung der neuen Viruspartikel durch Ausstülpung der Wirtszellmembran (7). Der Eintritt von Filoviren in die Wirtszelle ist ein hochkomplexer, mehrschrittiger Prozess. Über ihr Oberflächenglykoprotein (GP) heften sich die Viruspartikel an Haftstrukturen (Rezeptoren) auf der Wirtszelloberfläche und werden in einem Makropinozytose-ähnlichen Mechanismus endosomal in die Zelle aufgenommen (internalisiert) (Davey et al., 2017; Rhein and Maury, 2015). Aufgrund des breiten Wirtszelltropismus (Martines et al., 2015) ist davon auszugehen, dass Filoviren eine Vielzahl unterschiedlicher Wirtszellrezeptoren für ihren Eintritt in verschiedene Zellspezies nutzen. Für EBOV und MARV konnte unter anderem der T-Zell- und Mukosaepithelzell-Rezeptor TIM-1 (T-Zell-Immunglobulin und Mucin Domäne-1) als Eintrittsrezeptor identifiziert werden (Kondratowicz et al., 2011). Nachdem das in Endosomen verpackte Virus ins



**Abb. 7: Der Filovirus-Verbreitungszyklus.** (1) Viruspartikel heften über ihr Oberflächenprotein, das Glykoprotein GP (gelbe Kegel) an Haftstrukturen (Rezeptoren; orange) auf der Wirtszelloberfläche an und werden über Endozytose in die Zelle aufgenommen. (2) Hier verschmelzen die virale Virushülle und die zelluläre Membran miteinander, um das Filovirusnukleokapsid im Zytoplasma freizusetzen. (3) Nun transkribiert der Polymerasekomplex, bestehend aus der katalytischen Untereinheit L (hellgraue, ovale Scheiben), ihrem Kofaktor VP35 (dunkelgrüne Kreise) sowie dem Transkriptionsverstärker VP30 (schwarze Kreise), die filoviralen mRNAs (mit 5'-Kappe und Poly(A)-Schwanz), welche anschließend (4) mithilfe der Wirtszelltranslationsmaschinerie in die Filovirusproteine translatiert werden. Diese können nun genutzt werden, um neue Genomkopien in der Replikation (5) herzustellen. In der sekundären Transkription (5) werden die neuen Genome wiederum vielfach in neue mRNAs umgeschrieben, welche entsprechend in virale Proteine translatiert werden. Diese können anschließend zusammen mit neu gebildeten Genomkopien zu neuen Viruspartikeln zusammengebaut (6) und an der Plasmamembran ausgeknospt/freigesetzt werden (7). In Anlehnung an ICTV-Report, März 2019: [https://talk.ictvonline.org/ictv-reports/ictv\\_online\\_report/negative-sense-ma-viruses/mononegavirales/w/filoviridae](https://talk.ictvonline.org/ictv-reports/ictv_online_report/negative-sense-ma-viruses/mononegavirales/w/filoviridae).

Zytoplasma aufgenommen wurde, prozessieren Kathepsine das auf der Viruspartikeloberfläche befindliche GP<sub>1,2</sub>. Auf diese Weise wird die Rezeptorbindestelle von GP1 freigelegt, welche nun die Bindung von GP<sub>1,2</sub> an den endosomalen Rezeptor Niemann Pick C1 (NPC1) vermittelt (Davey et al., 2017; Sadewasser et al., 2019). Infolgedessen verschmilzt die Virushülle mit der zellulären (endosomalen) Membran und das Virusnukleokapsid kann in das Zytoplasma freigesetzt werden. Dort schreibt der Polymerasekomplex (L und VP35) das virale Genom in mRNAs um (Primärtranskription). Diese frühe/primäre Transkription erfolgt ausschließlich durch die vom Virus mitgebrachten Proteine (Hume and Mühlberger, 2019). Die daraus resultierenden mRNAs werden nun mithilfe der Wirtszelltranslationsmaschinerie in virale Proteine translatiert (Hume and Mühlberger, 2019). Die Translation erfolgt an freien Ribosomen im Zytoplasma, außer im Falle von GP, welches am Endoplasmatischen Retikulum (ER) und Golgi-Apparat hergestellt und posttranslational modifiziert wird (Glykosylierungen und Acetylierungen; Spaltung in GP<sub>1,2</sub>) (Funke et al., 1995; Gordon et al., 2019; Neumann et al., 2007; Volchkov et al., 1998). Die Neubildung von Virusproteinen ist die Grundvoraussetzung für die Replikation des viralen Genoms sowie für die sekundäre

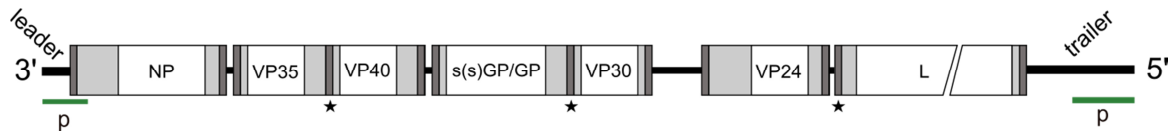


Transkription weiterer viraler mRNAs (Abb. 7; Kolesnikova et al., 2017). Replikation und Transkription nutzen beide die virale genomische RNA (vRNA) als Templat sowie den viralen Polymerasekomplex, der ausschließlich in 5'-3'-Richtung synthetisiert. Die Replikation des in 3'-5'-Richtung orientierten (-) RNA-Genoms muss folglich über einen Zwischenschritt erfolgen, in dem zunächst eine Intermediat-RNA gebildet wird, d.h., eine antigenomische, positiv-orientierte (+) RNA, auch copy RNA (cRNA) genannt. Diese dient dann im zweiten Schritt als Templat für die genomische vRNA (Mühlberger, 2007). Die Synthese viraler RNAs und auch der Zusammenbau neuer Nukleokapside erfolgt in sogenannten *inclusion bodies* (Einschlusskörperchen), die sich im Zytoplasma der infizierten Zelle ausbilden (Gordon et al., 2019; Hoenen et al., 2012). Mature Nukleokapside und matures GP<sub>1,2</sub> werden separat zur Plasmamembran transportiert, an der die Partikelauknoospung (*budding*) stattfindet. Der Transport von Nukleokapsiden ist Aktin-vermittelt und wird hauptsächlich von den drei viralen Proteinen NP, VP35 und VP24 gesteuert (Takamatsu et al., 2018). Das Matrixprotein VP40 assembliert über eine Dimerbildung zu Hexameren, während es ebenfalls zur Plasmamembran wandert (zumindest teilweise Aktin-vermittelt) und sich dort von innen an die Plasmamembran in geordneten Gitterstrukturen anlagert (Gordon et al., 2019; Madara et al., 2015; Takamatsu et al., 2018). Hier übernimmt es seine Hauptaufgabe, die Katalyse des Ausknospens, in dem es sich geschickt die zelluläre ESCRT-Maschinerie (Endosomal Sorting Complex Required for Transport) zunutze macht. Während dieses Prozesses wird GP<sub>1,2</sub> zum VP40-Hexamergitter an der Plasmamembran geführt und anschließend in die Zellmembran eingelagert. Diese umhüllt im letzten Schritt die herantransportierten Nukleokapside, woraufhin sich die Viruspartikel von der Zellmembran abschnüren können (Gordon et al., 2019; Madara et al., 2015).

## 1.7 Das EBOV-Genom

Filovirusgenome sind unsegmentierte, einzelsträngige RNA-Moleküle mit negativer Polarität (3'-5'-Ausrichtung), die eine Länge zwischen 15 kb (Thamnoviren) bis hin zu ca. 19 kb (Cuevaviren, Ebolaviren, Marburgviren) aufweisen und eine variable Zahl monocistronischer Gene in linearer Folge kodieren (Hume and Mühlberger, 2019). Die genomischen Sequenzen der Filovirusgenera unterscheiden sich um etwa 55 % voneinander (Bào et al., 2017). Die Genfolge im Genom ist grundsätzlich vergleichbar mit den meisten Mononegaviren: 3'-N-P-M-(G)-L-5' (alternative Terminologie für Filoviren: 3'-NP-VP35-VP40-(GP)-L-5'), unterscheidet sich jedoch dahingehend, dass zusätzliche Gene enthalten sind, die z.B. für den Filovirus-spezifischen Transkriptionsfaktor VP30 sowie das Nukleokapsidprotein VP24 codieren (im Falle der Genera *Cuevavirus*, *Ebolavirus* und *Marburgvirus*; Abb. 8) (Feldmann et al., 1992; Hume

and Mühlberger, 2019; Negredo et al., 2011; Sanchez et al., 1993). Eine Besonderheit der Ebolaviren und Cuevaviren ist, dass beide Filovirusgenera drei verschiedene Glykoproteine von ihrem GP-Gen exprimieren (Abb. 8); das strukturelle GP-Protein sowie zwei nicht-strukturelle GP-Proteine, soluble (sGP) und small soluble GP (ssGP), die durch ko-transkriptionelle Editierung der GP-mRNA entstehen (Mehedi et al., 2011; Negredo et al., 2011; Rougeron et al., 2015). Die sieben EBOV-Gene sind entweder



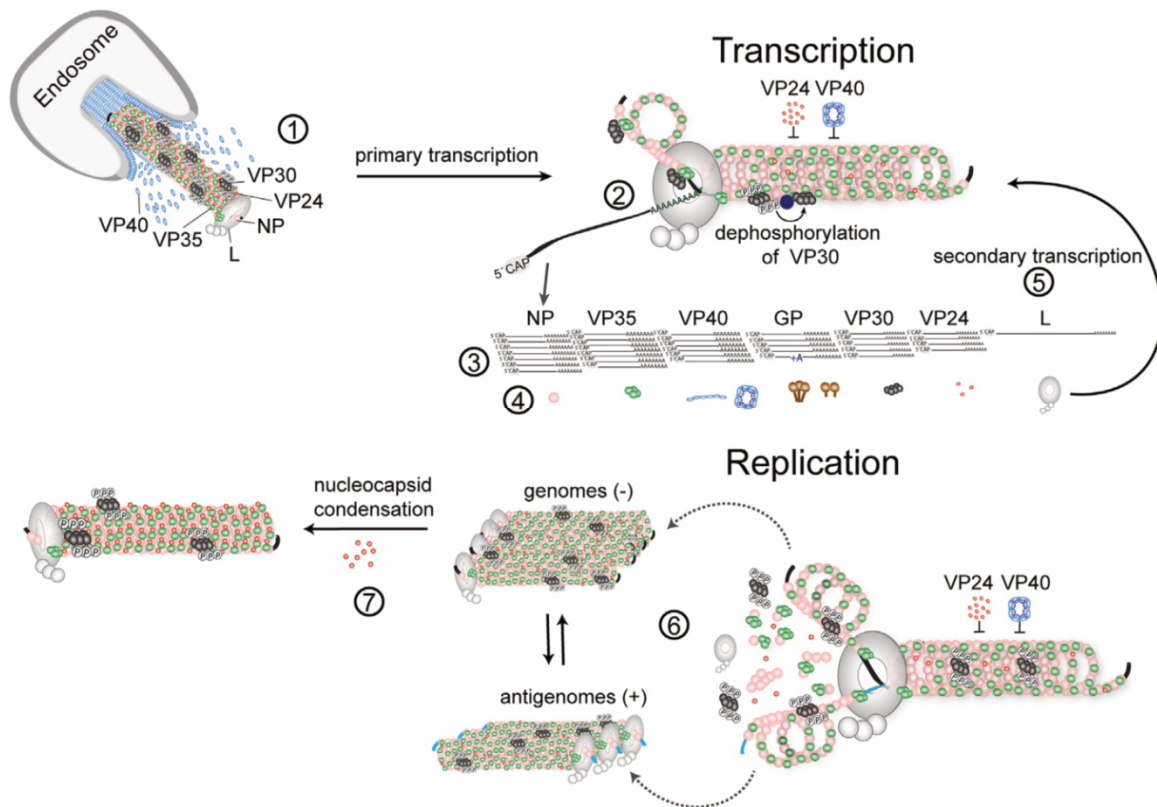
**Abb. 8: Das EBOV-Genom.** Die sieben offenen Leserahmen (ORF, weiße Kästen) des EBOV kodieren für sieben strukturelle (NP, VP35, VP40, GP, VP24, L) und zwei nicht-strukturelle (sGP, ssGP) Proteine. Sie werden von den 3'-Leader und 5'-Trailer-Sequenzen (schwarze, dicke Striche) flankiert, welche den genomischen oder antigenomischen Replikationspromotor (p; grüne Striche) sowie wichtige Enkapsidierungssignale enthalten. Das EBOV besitzt wie alle Filoviren lange UTRs (untranslatierte Regionen; hellgraue Balken), die u.a. die konservierten Genstart- (GS) und Genend (GE)-Signale (dunkelgraue Kästen) enthalten. Genstarts und -enden sind entweder durch intergenische Regionen (IGR) getrennt (dünne schwarze horizontale Balken) oder sie überlappen (Sternsymbol). Das VP24-Gen besitzt zwei GE-Signale, von denen das zweite mit dem GS-Signal des L-Gens überlappt.

durch nicht-konservierte intergenische Regionen (IGR) voneinander getrennt oder überlappen (Abb. 8; Brauburger et al., 2014, 2016). Die offenen Leserahmen (open reading frames, ORFs) werden von konservierten extragenischen 3'- und 5'-terminalen Sequenzen flankiert (= 3'-Leader und 5'-Trailer, siehe Abb. 8). Diese enthalten wichtige regulatorische Signale der Replikation und Transkription (siehe 1.8.1.2) sowie der Enkapsidierung (Calain et al., 1999; Hume and Mühlberger, 2019; Sanchez et al., 1993; Volchkov et al., 1999). Eine Besonderheit des EBOV-Genoms und der Filovireng Genome im Allgemeinen ist, dass sie im Vergleich zu anderen Viren der *Mononegavirales* stromaufwärts und -abwärts ihrer ORFs ausgesprochen lange nicht-kodierende Bereiche (untranslated regions, UTRs) aufweisen (siehe Abb. 8). Diese UTRs gelten als Hauptregulatoren der filoviralen Transkription (Brauburger et al., 2016; Neumann et al., 2009). U.a., weil sie die konservierten Genstart- (GS) und Genend (GE)-Signale enthalten, welche die Initiation und Termination der Transkription regulieren (siehe 1.8.1.2).

## 1.8 EBOV-Replikation und -Transkription

Ein gemeinsames Merkmal der NNS-Viren ist, dass ausschließlich enkapsidierte virale RNA (d.h. als Nukleokapsid verpackte RNA, siehe 1.5) als Templat der Replikation und Transkription akzeptiert wird, nicht jedoch nackte RNA (Conzelmann, 2004). *In vitro*-Experimente mit aufgereinigten viralen Polymerasen aus NNS-RNA-Viren zeigen jedoch, dass die entsprechenden viralen RNA-abhängigen RNA-Polymerasen (L + Kofaktor P; EBOV: VP35) prinzipiell in der Lage sind, auch nackte, nicht mit Nukleoprotein

(N; EBOV: NP) enkapsidierte RNA als Templat zu nutzen und zu replizieren (Morin et al., 2012, 2017; Olsen et al., 2018). Allerdings erlauben diese *in vitro*-Systeme ausschließlich die Synthese relativ kurzer RNA-Fragmente bis zu einer Länge von maximal 19-25 Nt. Aus experimentellen Daten zum Modellvirussystem VSV lässt sich ableiten, dass die Initiation der RNA-Synthese zwar theoretisch auf nackter RNA möglich ist, die effiziente Elongation der RNA sowie die Erkennung von mRNA-Terminationssignalen (GE-Signale) hingegen absolut abhängig vom gut verpackten N-RNA-Templat ist (Morin et al., 2012, 2017). Die Fähigkeit der NNS-Polymerasen, auch nackte RNA zu kopieren, ist konsistent mit der Annahme, dass das Nukleoprotein zumindest zeitweise während der Replikation und Transkription vom RNA-Templatstrang entfernt werden muss, damit die RNA in den Polymerasekomplex eingefädelt und abgelesen werden kann (Morin et al., 2017). Darauf verweisen auch Strukturdaten der VS(I)V-, RABV- und (H)RSV-Nukleoprotein:RNA-Komplexe (Albertini, 2006; Green et al., 2006; Tawar et al., 2009). Im Falle von VSV scheint das Entfernen von N vom RNA-Templat transient aufzutreten, da die RNA auch während der Polymerisation nicht zugänglich für Nukleasen ist. Auch die vielfachen N-N-Interaktionen in der N-RNA sprechen dagegen, dass die N-Moleküle vollständig vom RNA-Templat dissoziieren (Morin et al., 2017). Effiziente *in vitro*-Systeme für EBOV sind bislang nicht vorhanden. Es gibt lediglich einen Ansatz mit aus Insektenzellen aufgereinigter viraler Polymerase (L + VP35), der zeigt, dass die EBOV-Polymerase prinzipiell ebenfalls in der Lage ist, ein nacktes, 11 Nukleotide langes RNA-Templat zu kopieren (Tchesnokov et al., 2018). Für die Untersuchung der EBOV-Replikation und -Transkription werden vor allem sogenannte Minigenomsysteme verwendet (siehe 1.9). Diese Systeme haben den Vorteil, dass sie eine separate Betrachtung der viralen Replikation und Transkription unter Verwendung aller viralen Proteine oder nur eines Teils der viralen Proteine ermöglichen. Auf diese Weise konnte gezeigt werden, dass drei der insgesamt fünf Nukleokapsidproteine (siehe 1.5) ausreichend sind, um die virale Replikation zu ermöglichen. Dies sind zum einen der virale Polymerasekomplex aus L und VP35 und zum anderen das Nukleoprotein NP, das die virale genomische RNA enkapsidiert und somit das optimale Templat für den Polymerasekomplex bereitstellt (Mühlberger, 2007; Mühlberger et al., 1999). Die virale Transkription benötigt jedoch ein weiteres Nukleokapsidprotein, den Transkriptionsfaktor VP30 (Feldmann et al., 1992; Mühlberger et al., 1999; Sanchez and Kiley, 1987). Es ist davon auszugehen, dass Transkriptionskompetente Nukleokapside ohne das virale Protein 24 (VP24) vorliegen müssen, bzw. nur wenig VP24 assoziiert sein darf, da VP24 das Nukleokapsid stark kondensiert und das RNA-Templat schwer zugänglich für den Polymerasekomplex macht (Hoenen et al., 2019). Somit ergibt sich das nachfolgende hypothetische Modell der EBOV-Transkription



**Abb. 9: Modell der EBOV-Replikation und -Transkription.** Farbschema wie in Abb. 5 und Abb. 7 (Hume and Mühlberger, 2019).

und -Replikation [Abb. 9; (Hume and Mühlberger, 2019)]: Nachdem die EBOV-Nukleokapside aus den Endosomen in das Zytoplasma der infizierten Zelle freigesetzt wurden (1), kann die Primärtranskription erfolgen (2). In diesem ersten Transkriptionsschritt werden ausschließlich die im Nukleokapsid mitgebrachten viralen Proteine zur RNA-Synthese genutzt (Vgl. 1.6). Dabei bindet VP35, der Kofaktor des Polymerasekomplexes, an NP und dirigiert so die katalytische Polymeraseuntereinheit L zum NP-RNA-Komplex. Es ist bislang noch nicht vollständig geklärt, wie genau VP30 seine Funktion der Transkriptionssteigerung ausführt, d.h., wann, wo und ob VP30 in den Polymerasekomplex mitaufgenommen wird. Sicher ist jedoch, dass der Phosphorylierungsstatus von VP30 eine entscheidende Rolle in der Regulation der primären und sekundären Transkription spielt (Biedenkopf et al., 2013, 2016a). So muss VP30 als Hexamer und dephosphoryliert vorliegen, um seine Funktion als Transkriptionsaktivator auszuüben (siehe 1.8.2). Die EBOV-Transkription nutzt einen sequentiellen Stopp-Start-Mechanismus, der zur Bildung von sieben (zumeist) monocistronischen (ein Gen-) mRNAs führt. Hier erkennt der Polymerasekomplex konservierte Genstart (GS)- und Genend (GE)-Signale (siehe 1.8.1.1) in der genomischen RNA, an denen entweder die mRNA-Synthese initiiert oder terminiert wird. Grundsätzlich ist die sequentielle Transkription ein hochregulierter Prozess. So kann die



Polymerase an internen Genen nur dann am GS-Signal initiieren, sofern die zuvor transkribierte mRNA erfolgreich terminiert wurde (siehe 1.8.1.1). Man geht davon aus, dass die virale Polymerase an jedem GE-Signal entweder zum nächsten GS-Signal weiterscannt und reinitiiert oder von Zeit zu Zeit vom Templat dissoziiert und dann am 3'-Ende des Genoms mit einer neuen Polymerisationsrunde beginnt (Brauburger et al., 2016). Dieser Stopp-Start-Transkriptionsmodus und die Annahme, dass die virale Transkription immer vom 3'-Terminus aus initiiert werden muss, erklären den zu beobachtenden mRNA-Gradienten vom ersten NP-Gen bis hin zum letzten L-Gen (3). Konkret heißt dies, dass Gene, die näher am 3'-Terminus liegen, häufiger abgelesen werden als die 5'-terminalen Gene (Mühlberger, 2007; Shabman et al., 2014). Neue Sequenzierdaten weisen allerdings darauf hin, dass die Abnahmen der mRNA-Spiegel nicht ganz gleichmäßig verlaufen. Vielmehr scheinen die mRNAs der ersten vier Gene (NP, VP35, VP40, GP) in ungefähr vergleichbaren Mengen exprimiert zu sein, während die VP30- und VP24-mRNAs mittlere Expressionsraten aufweisen und die L-mRNA am geringsten exprimiert ist. Darüber hinaus scheinen sich insbesondere die Levels der NP- und der GP-mRNA im Verlauf der Infektion zu verändern. Während die Expression der GP-mRNA steigt, sinkt die NP-mRNA-Expression (Albariño et al., 2018). Noch während der Transkription der jeweiligen mRNAs überträgt eine Domäne der Polymeraseuntereinheit L die 5'-Kappe und methyliert diese. Der Poly(A)-Schwanz der naszierenden mRNA entsteht durch ein Stottern der viralen Polymerase an kurzen Uridin (U)-Trakten der GE-Signale (siehe 1.8.1.1). Ein solches Stottern der Polymerase führt, ebenfalls kotranskriptionell, zur Insertion eines oder mehrerer zusätzlicher Adenosine in die GP-mRNA [= RNA-Editierung; in Abb. 9 mit +A gekennzeichnet; zusammengefasst in (Hume and Mühlberger, 2019)]. Die in der Primärtranskription gebildeten mRNAs ähneln mit ihrer 5'-Kappe und ihrem 3'-Poly(A)-Schwanz den zellulären mRNAs und werden somit von der zellulären Translationsmaschinerie erkannt und in Proteine translatiert (4). Diese werden zeitgleich für die virale Replikation (6) und die sekundäre Transkription (5) benötigt. Beide Prozesse werden durch neu gebildete VP40- und VP24-Proteine negativ reguliert, bzw. in Schach gehalten (Hoenen et al., 2010). VP35 dient vermutlich als Chaperon, das ungebundenes NP vor der Aggregation bewahrt, bis es neu synthetisierte Genome oder Antigenome enkapsidieren kann (Kirchdoerfer et al., 2015). Während dephosphoryliertes VP30 die sekundäre Transkription begünstigt, unterdrückt es gleichzeitig die virale Replikation (Biedenkopf et al., 2013). Daher muss VP30 für eine effiziente Replikation durch zelluläre Phosphatasen wieder phosphoryliert werden (1.8.2; Takamatsu et al., 2019). Bei der Genomreplikation wird zunächst ein positiv orientiertes Antigenom gebildet, welches anschließend als Templat für die Synthese genomischer RNA dient. Sowohl das Genom

als auch das Antigenom müssen für eine effiziente Replikation von NP enkapsidiert vorliegen (Brauburger et al., 2014b). Allerdings wird ausschließlich das virale Genom in neue Viruspartikel verpackt. Das Antigenom ist ein reines Replikationsintermediat. Die neu gebildeten Genome und Antigenome werden von VP24, dem Nukleokapsidmaturierungsfaktor, kondensiert (7). Auf diese Weise verschiebt sich der Replikationszyklus in Richtung Assemblierung. Genomreplikation (und Sekundärtranskription) sowie die Assemblierung neuer Nukleokapside finden in zytosolischen Einschlusskörperchen statt (Brauburger et al., 2014b; Hoenen et al., 2012). Über die Bindung an NP wird phosphoryliertes VP30 mit in die neuen Nukleokapside verpackt, welche anschließend für die Ausknospung zur Zellmembran transportiert werden können (Banadyga et al., 2017; Dolnik et al., 2015; Hoenen et al., 2012).

## 1.8.1 Regulatorische Elemente der EBOV-Replikation und -Transkription

### 1.8.1.1 Funktionen der Genstart- und Genendsignale

Jedes der EBOV-Gene wird von hochkonservierten Genstart (GS)- und Genend (GE)-Signalen flankiert (Neumann et al., 2009). Die 12 Nukleotide umfassenden GS-Signale weichen lediglich in ihrer 3. Position in zwei Fällen ab (NP und L; Cytosin statt Adenosin; Tab. 1). Die biologische Relevanz dessen für die Transkription bleibt noch zu klären.

	GS-Signal (3' → 5')	GE-Signal (3' → 5')
Konsensussequenz	CUC <sub>C/A</sub> CUUCUAAUU	UAAUUC <sub>L/(U)</sub> <sub>5/6</sub>
NP	CUC <u>C</u> CUUCUAAUU	UAAUUCUUUUUU <u>U</u>
VP35	CUACUUCUAAUU	UAAUUCUUUUUU
VP40	CUACUUCUAAUU	UAAUUCUUUUUU <u>U</u>
GP	CUACUUCUAAUU	UAAUUCUUUUUU
VP30	CUACUUCUAAUU	UAAUUCUUUUUU
VP24	CUACUUCUAAUU	UAAUUCUUUUUU <u>U</u>
L	CUC <u>C</u> CUUCUAAUU	UAAUU <u>U</u> UUUUUU <u>U</u>

**Tab. 1: EBOV-Genstart-(GS) und Genend (GE)-Signale der genomischen RNA und ihre Konsensussequenzen.** Dick gedruckte und unterstrichene Nukleotide geben Unterschiede in der ansonsten konservierten Sequenz an. In der Konsensussequenz sind die konservierten Nukleotide dick gedruckt. Nukleotidvariationen an nicht-konservierten Positionen sind durch einen Schrägstrich gekennzeichnet.

Die virale Polymerase betritt das NP-RNA-Templat mutmaßlich immer an dem am 3'-Genomende vermuteten Replikations- oder Transkriptionspromotor (siehe 1.8.1.2),

bevor sie zum ersten GS-Signal (= Transkriptionsstartsequenz, TSS) des NP-Gens scannt, um dessen Transkription zu initiieren (Hume and Mühlberger, 2019). Die Konsensussequenz der GS-Signale der Ebola-, Cueva- und Marburgviren lautet 3'-**CUNCUUNUAAUU**-5' (N = variables Nukleotid). Ebolaviren haben an Position 7 eine Präferenz für Cytidin (C), wohingegen bei Marburg- und Cuevaviren an dieser Stelle Purine (entweder A oder G, Guanosin) zu finden sind (Hume and Mühlberger, 2019). Das erste Nukleotid aller GS-Signale ist immer ein Cytidin (C, Tab. 1). Entsprechend beginnt jede der mRNAs mit einem Guanosin. Dieses erste G-Nukleotid scheint wichtig für das Anbringen der 5'-Kappe zu sein (Martin et al., 2018). Ähnlich zu VSV-mRNAs besitzen EBOV-mRNAs eine methylierte 5'-<sup>m7</sup>Gppp-Kappe (Martin et al., 2018; Weik et al., 2002). Für Sudan-EBOV (SUDV) konnte gezeigt werden, dass die neu synthetisierten mRNAs wie bei VSV 31 Nt nach ihrer Initiation ihre 5'-Kappe erhalten (Martin et al., 2018). Nach der Transkriptionsinitiation und dem Anbringen der 5'-Kappe bewegt sich die Polymerase entlang des Templatstrangs und elongiert die mRNA so lange, bis sie das erste GE-Signal erreicht. Auch die GE-Signale sind hochkonserviert für alle Filovirusgenera (Tab. 1). Sie enden obligatorisch mit einem 5-6 Nt langen U-Trakt, welcher ein Stottern der Polymerase auslöst, infolgedessen ein Poly(A)-Schwanz von höchstwahrscheinlich 100-300 Nt angehängt wird (Brauburger et al., 2014b; Whelan et al., 2004). Nach der Transkriptionstermination scannt die virale Polymerase bis zum nächsten GS-Signal. Das GE-Signal des folgenden Gens kann vom nachfolgenden GS-

Gengrenze	Anordnung der GS- und GE-Signale
NP/VP35	<b>UAAUUCUUUUU</b> GAUUA <b>CUACUUCUAAUU</b>
VP35/VP40	<b>CUACUUCUAAUU</b> <b>CUUUUU</b>
VP40/GP	<b>UAAUUCUUUUU</b> AGCCU <b>CUACUUCUAAUU</b>
GP/VP30	<b>CUACUUCUAAUU</b> <b>CUUUUU</b>
VP30/VP24	<b>UAAUUCUUUUU</b> 144 Nt <b>CUACUUCUAAUU</b>
VP24/L	<b>UAAUUCUUUUU</b> CGGA <b>CUCCUUCUAAUU</b> <b>CUUUUU</b>

**Tab. 2: EBOV-Gengrenzen.** EBOV-Genstart- (GS; cyan) und Genend (GE; rot)-Signale sind entweder durch variable, nicht-konservierte intergenische Regionen (IGR; schwarz) unterschiedlicher Länge getrennt oder überlappen (Unterstrichene, schwarze Nt). Unterstrichene, farbige Nt verweisen auf Unterschiede in der ansonsten konservierten Sequenz.

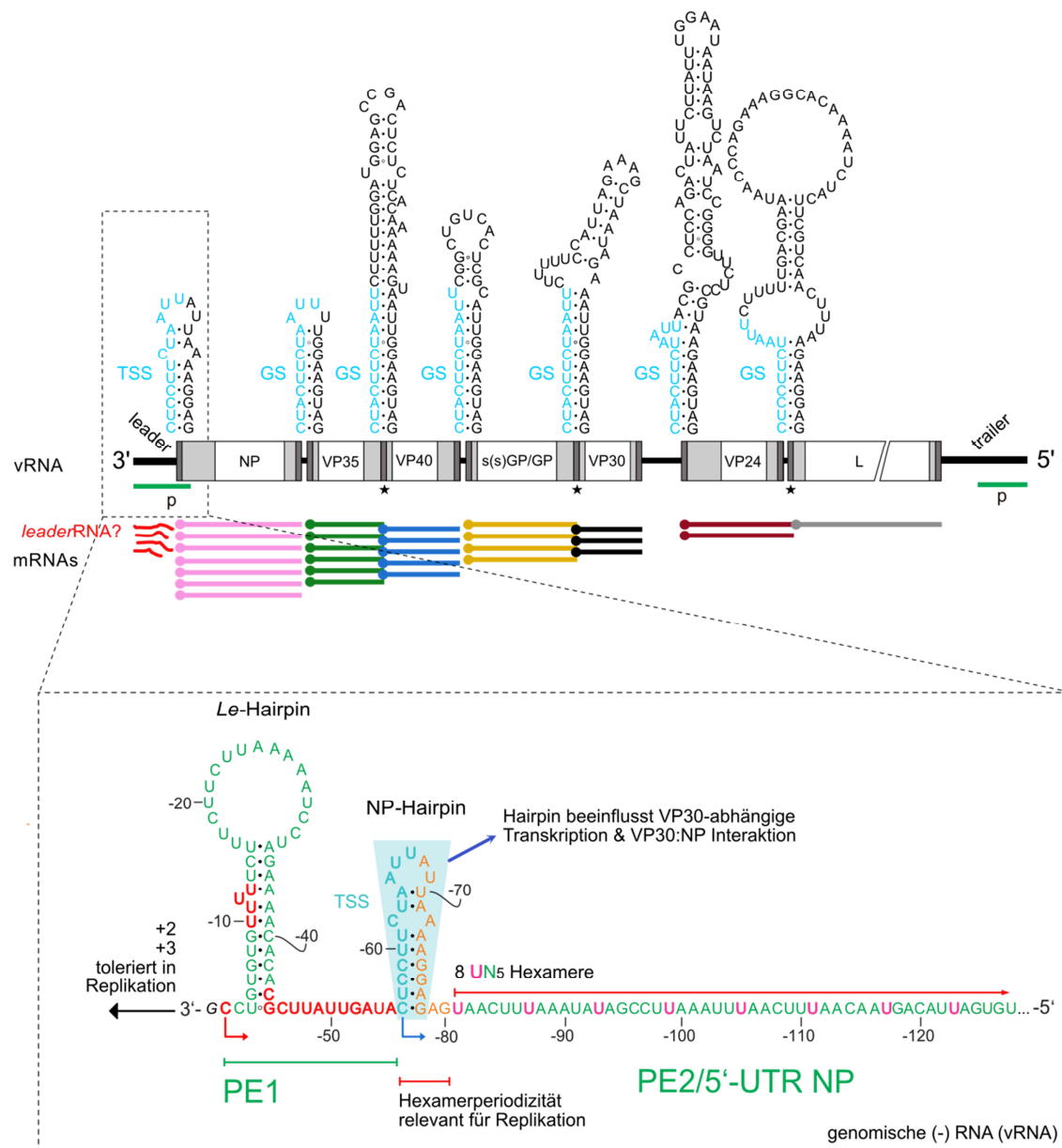
Signal durch eine kurze (4-5 Nt) oder lange (144 Nt) intergenische Region (IGR) getrennt sein oder beide Signale überlappen und teilen sich ein konserviertes 3'-**UAAUU**-5'-Pentamer (Tab. 2). Die überlappenden GS-Signale zwingen die virale Polymerase ein

kurzes Stück rückwärts zu scannen, um am GS-Signal die Synthese der nächsten mRNA zu initiieren. Es konnte gezeigt werden, dass die Transkriptionsinitiation an Genüberlappungen genauso effizient erfolgt wie bei IGR-separierten GS- und GE-Signalen (Brauburger et al., 2016). Grundvoraussetzung für die effiziente Transkriptionsinitiation an internen Genen ist die erfolgreiche Termination der zuvor transkribierten mRNA. Wenn die virale Polymerase GS- und GE-Signale fälschlicherweise „überliest“, entstehen zu einem kleineren Prozentsatz auch sogenannte Readthrough-mRNAs, d.h., mRNAs, deren Sequenz über die Gengrenzen hinausgeht (Brauburger et al., 2014a; Shabman et al., 2014).

Interessanterweise sind laut bioinformatischer Vorhersage alle filoviralen GS-Signale auf genomischer (siehe 1.8.1.2, Abb. 10) und korrespondierender mRNA/antigenomischer Ebene Teil von RNA-Sekundärstrukturen, sogenannten Hairpin- (Haarnadel) oder Stem-loop (Stamm-Schleife)-Strukturen (Mühlberger et al., 1996; Sanchez et al., 1993). Eine putative Funktion dieser Haarnadelstrukturen auf mRNA-Ebene als interne ribosomale Eintrittsstellen (internal ribosomal entry sites, IRES) konnte ausgeschlossen werden (Shabman et al., 2013; Weik et al., 2002). Für die Hairpinstruktur am ersten NP-Genstart (NP-Hairpin) wird postuliert, dass sie die Abhängigkeit der viralen Transkription vom Transkriptionsfaktor VP30 beeinflusst (Weik et al., 2002). Während die virale Transkription im Falle des nativen NP-Hairpins strikt VP30-abhängig erfolgt, scheint eine Destabilisierung des NP-Hairpins, vor allem auf mRNA-Ebene, die virale Transkription zu einem gewissen Ausmaß auch in Abwesenheit von VP30 zu ermöglichen. Bezugnehmend auf die biologische Situation bedeutet dies, dass VP30 möglicherweise dem Polymerasekomplex dabei hilft, Sekundärstrukturen in der Templatsequenz zu überwinden. Darüber hinaus wird im Falle der NP- und L-GS-Hairpins vermutet, dass sie Erkennungssignale für die Interaktion mit zellulären Typ 1 Topoisomerasen aufweisen. Takahashi *et al.* postulieren, dass die DNA-Topoisomerase I direkt mit EBOV-L interagiert. Sie vermuten, dass die virale Transkription und Replikation auf diese Weise gesteigert werden (Takahashi et al., 2013).

### **1.8.1.2 Die EBOV-Replikations- und Transkriptionspromotoren**

Der genomische und antigenomische EBOV-Replikationspromotor (siehe Abb. 10) befindet sich am 3'-terminalen Ende (Genompromotor ca. Nt -1 bis -128; Antigenompromotor ca. Nt 1 bis 176) der jeweiligen RNA (Calain and Roux, 1995; Volchkov et al., 1999; Weik et al., 2005). Eine Besonderheit des filoviralen Replikationspromotors im Gegensatz zu vielen anderen NNS-Viruspromotoren ist, dass er zweigeteilt (bipartite) ist (Weik et al., 2005). Nur Paramyxoviren wie das SeV nutzen ebenfalls zweigeteilte Promotoren (Calain and Roux, 1993; Kolakofsky, 2016; Le Mercier



**Abb. 10: Ebola-Virus-Genom und Zoom in den 3'-terminalen Replikations- und Transkriptionspromotorbereich.** Oben: Die 7 EBOV-Genstart (GS)-Signale (cyan) sind ein Teil von putativen RNA-Hairpinstrukturen. Watson-Crick-Basenpaarungen sind als schwarze, ausgefüllte Kreise dargestellt, G:U-Basenpaare durch ungefüllte Kreise. Farbschema wie in Abb. 8. Unterhalb des schematischen Genoms ist der zu beobachtende mRNA-Gradient angedeutet. In rot sind die putativen *leader* RNAs als abortive Vorlauf-Transkriptionsprodukte abgebildet. Unten: Der EBOV-Replikationspromotor (grün) ist zweigeteilt. Ein erstes Promotorelement (PE1) befindet sich zwischen Nt -1 und -55. Der Sequenzerhalt der rot markierten, dick gedruckten Nt ist obligat. PE2 ist charakterisiert durch 8 aufeinanderfolgende UN<sub>5</sub>-Hexamere zwischen Nt -81 und -128 (Die initialen Uridine der Hexamere sind pink eingefärbt). Drei dieser UN<sub>5</sub>-Hexamere sind essenziell für eine residuale Replikationsaktivität. PE1 und 2 werden durch einen sogenannten Spacer-Bereich (Nt -56 bis -80; orange) getrennt, der das erste NP-GS-Signal (= Transkriptionsstartsequenz, TSS; cyan) enthält. Der Spacer folgt der „Rule of 6“, weshalb ausschließlich Insertionen und Deletionen eines Vielfachen von 6 (Hexamerperiodizität) replikationsaktiv sind. Die Antigenomsynthese wird am zweiten Genomnukleotid initiiert (roter Pfeil), unabhängig von der Basenidentität der ersten Genomnukleotids oder Extensionen (+2 und +3) des 3'-Genomendes.

and Kolakofsky, 2019; Tapparel et al., 1998). Das erste Promotorelement (PE1) des genomischen Replikationspromotors ist in der 3'-Leader-Region lokalisiert, wohingegen das zweite Promotorelement (PE2, Pos. -81 bis -128), ein konsekutives 3'-UNNNNN-5'-

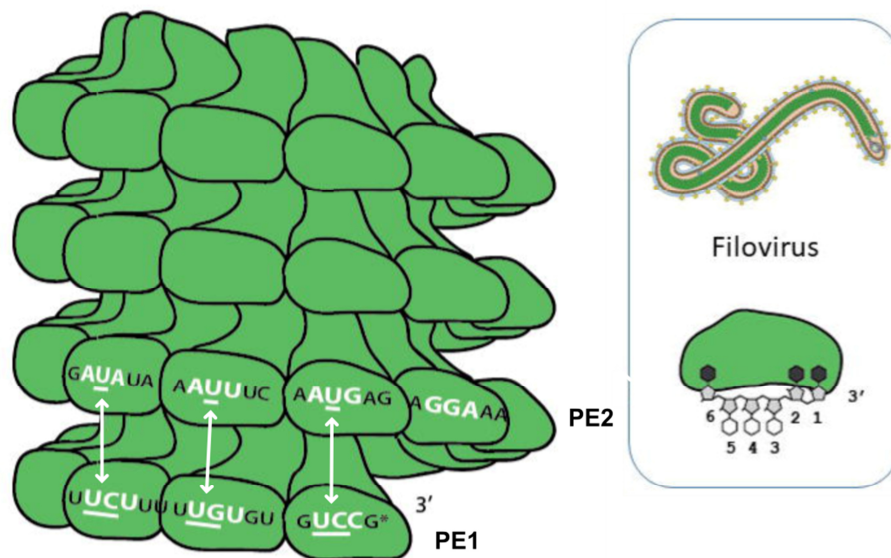
Hexamerraster, weiter intern lokalisiert ist (Abb. 10). Mindestens drei der UN<sub>5</sub>-Hexamere in PE2 müssen intakt sein, um residuale Replikation zu ermöglichen (Weik et al., 2005). Mutationsexperimente zeigten, dass der Sequenzerhalt an den Nukleotidpositionen -10 bis -13 sowie -44 bis -55 unerlässlich für die virale Replikation ist (Weik et al., 2005). Eine Deletion der terminalen 25 Nt des Trailer-Promotors eliminiert die Neusynthese genomischer viraler RNA (vRNA) und beschränkt die Replikation auf die Synthese des antigenomischen Replikationsintermediats (Deflubé et al., 2019; Olsen et al., 2018). Daraus lässt sich schließen, dass diese Nukleotide ein wesentlicher Bestandteil des Replikationspromotors sind. Darüber hinaus kann aber ohne die terminalen 55 Trailer-Nukleotide kein Volllängenvirus (Virus rescue) aus Minigenomsystemen isoliert werden. Dies unterstreicht die Wichtigkeit dieser Nukleotide für die virale Replikation und/oder Enkapsidierung (Hoenen et al., 2010; Sztuba-Solinska et al., 2016). Auch die hohe Sequenzhomologität der terminalen 50 Nt des 3'-Leaders und 5'-Trailers spricht für die Relevanz dieser Nukleotide (Schlereth et al., 2016). Eine aktuelle Studie zeigt, dass die Synthese des Antigenoms nicht am äußersten 3'-Ende des Genoms initiiert wird, sondern an Position 2 (Deflubé et al., 2019). Eine Verlängerung des Genoms um zwei bis drei Nukleotide verändert die Initiationsposition der Antigenomsynthese nicht (Abb. 10). Somit ist das erste Genomnukleotid, welches auch nicht im 3'-terminalen Leader und Trailer konserviert ist, nicht Teil des Antigenomtemplats für die vRNA-Synthese und muss nachträglich angefügt werden. Die 45 3'-terminalen Genomnukleotide sind ebenfalls in der Lage, eine Stem-loop-Struktur (Leader (Le)-Hairpin, Abb. 10) auszubilden. Unter Verwendung nackter RNA wurde sowohl die Existenz des NP-Hairpins als auch des Leader-Hairpins zumindest *in vitro* experimentell validiert (Schlereth et al., 2016; Weik et al., 2002, 2005). Theoretisch denkbar wäre eine Rolle des Leader-Hairpins während der viralen Replikation, um in einer Art *Backpriming*-Mechanismus das nicht im Antigenomtemplat enthaltene erste 3'-terminale Genomnukleotid zu ergänzen (Deflubé et al., 2019). Alle bisherigen Daten deuten jedoch darauf hin, dass vielmehr die Leadersequenz als seine Sekundärstruktur eine Rolle in der Replikation spielt (Crary et al., 2003; Weik et al., 2005). Der größte Unterschied zwischen dem genomischen und antigenomischen Replikationspromotor ist, dass der Antigenompromoter lediglich die vRNA-Synthese steuert, während der Genomreplikationspromotor überlappende Bereiche mit dem putativen Transkriptionspromotor aufweist und somit Replikation und Transkription in Konkurrenz treten. Es wird vermutet, dass der Transkriptionspromotor 3' vom ersten GS-Signal lokalisiert ist. Allerdings konnte der Transkriptionspromotor bislang auf keine bestimmten Nukleotide des Leaders eingegrenzt werden (Hume and Mühlberger, 2019). Die mutmaßliche Transkriptionsinitiation am 3'-Ende des Genoms lässt die Annahme zu,

dass vor der Synthese der ersten NP-mRNA ein kurzes Leader-Transkript (*leaderRNA*) synthetisiert wird, welches jedoch abgebrochen wird, sobald das erste GS-Signal vom Polymerasekomplex erkannt wird (Abb. 10). Solche Leader-Transkripte konnten für einige der NNS-Viren mit sequentieller Stopp-Start-Transkription nachgewiesen werden (Noton and Fearn, 2015; Ogino and Green, 2019; Whelan et al., 2004). Bislang gibt es in der Literatur noch keine bestätigenden Hinweise auf die Existenz filoviraler *leaderRNAs*. Aufgrund der Konkurrenz von Replikations- und Transkriptionspromotor am 3'-Genomende wird vermutet, dass der Antigenomreplikationspromotor stärker ist als der Genomreplikationspromotor, ähnlich wie es bei anderen NNS-RNA-Viren der Fall ist (Le Mercier et al., 2003; McGivern et al., 2005; Vulli  moz et al., 2005). Infolgedessen wird mehr genomische vRNA synthetisiert als antigenomische cRNA. Vergleichende Studien zur Replikationspromotorst  rke in Filoviren gibt es jedoch nicht. Allerdings werden Transkriptions-defiziente MARV-Minigenome st  rker repliziert als Transkriptionskompetente Minigenome, was die These der konkurrierenden Replikation und Transkription am Genomreplikationspromotor unterst  tzt (Enterlein et al., 2009). Es ist noch ungekl  rt,   ber welchen Mechanismus die virale Polymerase w  hrend der Antigenomsynthese alle transkriptionsrelevanten Signale am ersten Gen und an internen Gen  berg  ngen ignoriert und anstelle von einzelnen mRNAs ein Volll  ngenantigenom synthetisiert.

#### **1.8.1.3 Die Relevanz der „Rule of 6“ f  r die EBOV-Replikation**

Die Genom-Replikationspromotorelemente PE1 und PE2 sind durch einen Spacer-Bereich getrennt. Dieser enth  lt den NP-Hairpin sowie zwei zus  tzliche Nt und umfasst folglich insgesamt 25 Nt (Nt -56 bis -80; Abb. 10). Insertionen oder Deletionen im Bereich des Spacers werden nur dann von der viralen Polymerase toleriert, wenn sie ein Vielfaches der Zahl 6 umfassen (Weik et al., 2005). Somit unterliegt der Spacer der „Rule of 6“. Diese ist auch relevant f  r Paramyxoviren, welche ebenfalls zweigeteilte Replikationspromotoren aufweisen. Allerdings umfasst die „Rule of 6“ bei der Paramyxovirusreplikation nicht ausschlie  lich den Promotorbereich, sondern bezieht sich auf die Gesamtl  nge des Genoms. Filovirusgenoml  ngen sind per se kein Vielfaches von 6. Nichtsdestoweniger gibt es auch bei den Paramyxoviren einige Vertreter, bei denen die Konformit  t mit der „Rule of 6“ nicht strikt erf  llt sein muss (Calain and Roux, 1993; Kolakofsky et al., 1998; Matsumoto et al., 2018; Murphy and Parks, 1997; Vulli  moz and Roux, 2001). Die obligate Hexamerperiodizit  t des Spacers wird mit der NP-Enkapsidierung in Zusammenhang gebracht (Tapparel et al., 1998; Vulli  moz and Roux, 2001; Weik et al., 2005). In Paramyxoviren- und Filovirennukleokapsiden sind jeweils ~ 13 N/NP-Molek  le pro Helixwindung auf dem

RNA-Templat assembliert, wobei pro NP-Molekül genau 6 Nt der (anti-)genomischen RNA gebunden werden (siehe Abb. 6; Egelman et al., 1989; Kirchdoerfer et al., 2019; Sugita et al., 2018; Wan et al., 2017). Dies sorgt dafür, dass PE1 und 2 auf derselben Seite der Nukleokapsidhelix liegen und somit vom Polymerasekomplex während der Replikationsinitiation möglicherweise simultan kontaktiert werden können (Abb. 11; Le Mercier and Kolakofsky, 2019; Matsumoto et al., 2018).



**Abb. 11: Anordnung von den Replikationspromotorelementen PE1 und PE2 im helikalen EBOV-Nukleokapsid.** Weiße Pfeile verweisen auf die Möglichkeit, dass PE1 sowie mind. drei der UN<sub>5</sub>-Hexamere in PE2 simultan vom Polymerasekomplex kontaktiert werden können, da sie auf derselben Seite der Helix angeordnet sind. Grundlage ist, dass 13 NP-Moleküle pro Helixwindung an jeweils 6 Nt der viralen RNA binden (umrahmtes Schema rechts). Adaptierte Abbildung nach (Le Mercier and Kolakofsky, 2019).

### 1.8.2 Die Rolle von VP30 in der EBOV-Transkriptionsinitiation

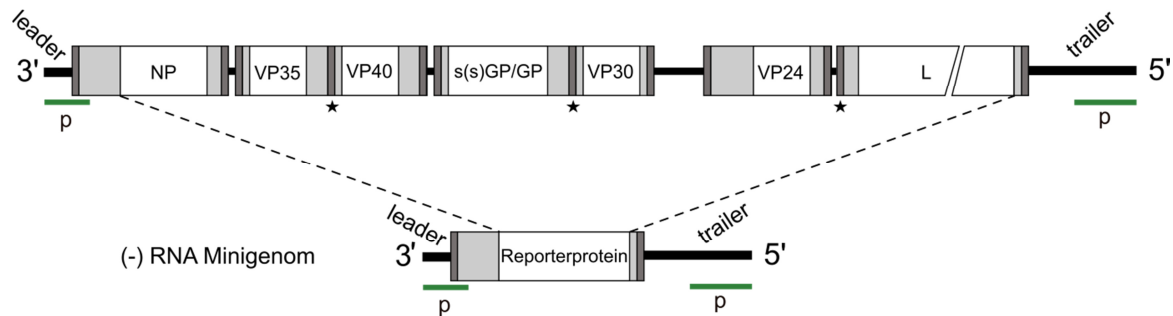
Neben den Polymerasekomplexkomponenten L und VP35 benötigt die Transkription des NP-RNA-Templats zusätzlich den Transkriptionsaktivator VP30. VP30 ist ein RNA-Bindeprotein, welches Hexamere ausbildet und mit NP, VP35 und L interagiert (Hume and Mühlberger, 2019; Schlereth et al., 2016). Eine *in vitro*-Studie zeigt, dass VP30 präferentiell an einzelsträngige RNA in der Nähe von Stem-loop-Strukturen bindet, jedoch ohne erkennbare Basenspezifität. Allerdings scheint eine gemischte Basenkomposition präferiert zu sein (Schlereth et al., 2016). Die Interaktion zwischen VP30 und VP35 ist RNA-vermittelt und dient möglicherweise dazu, den Polymerasekomplex zu stabilisieren und vor Dissoziation an den GS-Hairpinstrukturen (siehe Abb. 10) zu bewahren (Biedenkopf et al., 2016b). Möglicherweise ist VP30 auch an dem Wechsel von Transkription zu Replikation beteiligt (Biedenkopf et al., 2013; Martinez et al., 2011). Außerdem wird VP30 eine Rolle in der GP-mRNA-Editierung



zugeschrieben, wo es ähnlich wie an den GS-Signalen die Überwindung einer RNA-Sekundärstruktur ermöglicht und so das Stottern der viralen Polymerase unterdrücken soll (Mehedi et al., 2013). Die transkriptionsaktivierende Funktion von VP30 beruht maßgeblich auf seiner Phosphorylierung. Die endosomal in das Zytoplasma eingebrachten EBOV-Nukleokapside enthalten vor allem phosphorylierte VP30-Hexamere (Brauburger et al., 2014b). Um ihre Transkriptionsaktivatorfunktion ausüben zu können, müssen die VP30-Hexamere jedoch dephosphoryliert oder schwach phosphoryliert vorliegen (Biedenkopf et al., 2013; Modrof et al., 2003). Die Dephosphorylierung erfolgt durch zelluläre Proteinphosphatasen (PP), u.a. PP1 und PP2A (Ilinykh et al., 2014; Kruse et al., 2018). Die Dephosphorylierung durch PP2A wird über NP vermittelt, welches VP30 und PP2A simultan in virale Einschlusskörper (*inclusion bodies*) rekrutiert (Lier et al., 2017). Dephosphoryliertes VP30 hat eine verringerte Bindungsaffinität zu NP, jedoch eine höhere Bindungsaffinität zu VP35, den Kofaktor des Polymerasekomplexes. Möglicherweise kann auf diese Weise VP30 mit in den Polymerasekomplex aufgenommen werden und ihn so in einen optimierten Transkriptasekomplex überführen (Biedenkopf et al., 2016b; Takamatsu et al., 2019). Umkehrt führt eine Phosphorylierung von VP30, durch zelluläre Kinasen (u.a. Serine-Arginin Proteinkinase 1, SRPK1) vermittelt, zu einer verringerten Bindung an RNA und VP35, jedoch zeitgleich zu einer verstärkten NP-Bindung. Dies wiederum begünstigt seine Freisetzung aus dem Polymerasekomplex und damit die VP30-unabhängige Replikation. Aus diesen Daten lässt sich ableiten, dass eine dynamische Phosphorylierung von VP30 unerlässlich für den viralen Lebenszyklus ist, d.h. die Regulation von primärer und sekundärer Transkription sowie der Replikation (Biedenkopf et al., 2013, 2016b, 2016a; Takamatsu et al., 2019). Die Interaktion von VP30 und NP ist besonders engmaschig reguliert. Zu hohe Bindungsaffinitäten scheinen unerwünscht zu sein, weil sie mutmaßlich die oben genannte erforderliche Dynamik der Transkriptions- und Replikationsregulation stören. Niedrige Bindungsaffinitäten sind generell förderlich für die Transkriptionsaktivität. Allerdings gibt es Hinweise, dass die Interaktion von VP30 und NP während der initialen Phase der Transkription für die Überwindung der NP-GS-Hairpinstruktur wichtig ist (Kirchdoerfer et al., 2016; Xu et al., 2017).

## 1.9 EBOV-Minigenomsysteme

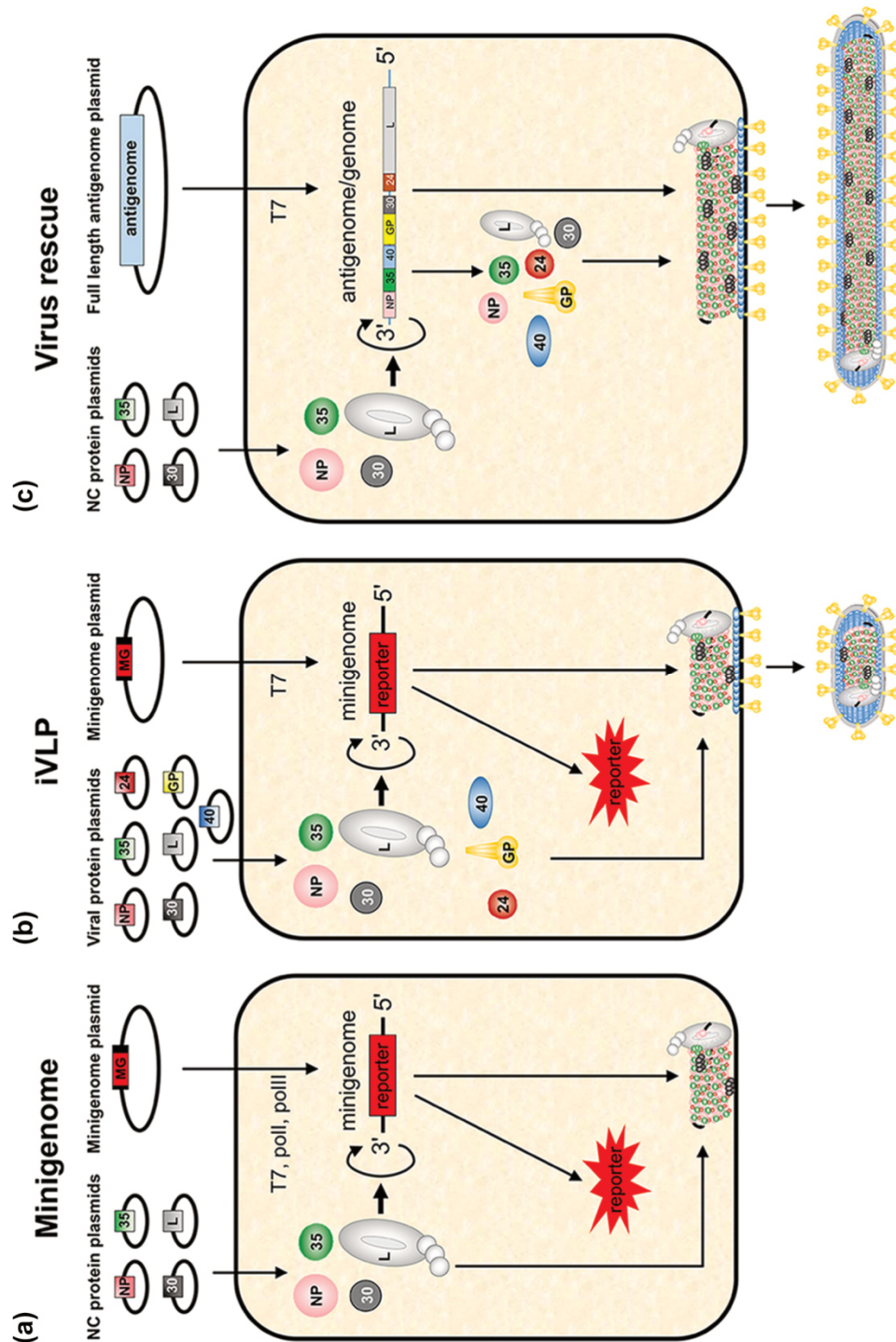
Mittels sogenannter Reverser Genetik werden Mutationen in das Virusgenom eingebracht und anschließend der dadurch hervorgerufene Phänotyp charakterisiert. Dies macht man sich in Transfektions-basierten Minigenomsystemen zunutze, um gezielt die Transkriptions- und/oder Replikationsrelevanz bestimmter RNA-Elemente oder Virusproteine zu analysieren. Zudem ermöglichen diese Systeme die Arbeit mit für den Menschen hochpathogenen Viren auf Biosicherheitslevel 1 oder 2-



**Abb. 12: Monocistronisches EBOV-Minigenom.** Farbschema wie in Abb. 8. Die offenen Leseraster (ORF) der EBOV-Proteine sind im Minigenom durch ein Reporterprotein, in der Regel eine Luziferase, ersetzt. Der regulatorische 3'-Leader und 5'-Trailer, welche die Replikations- und ggf. Transkriptionspromotoren enthalten, bleiben erhalten, ebenso wie die NP-5'-UTR (untranslatierte Region) mit dem ersten Genstart (GS)-Signal sowie die L-3'-UTR mit dem L-Genend (GE)-Signal.

(biosafety level 1, BSL-1) anstelle von BSL-4-Niveau (Hoenen et al., 2011, 2014). In der vorliegenden Arbeit wurden zumeist monocistronische (Abb. 12), d.h. aus einem ORF bestehende, oder bicistronische (Zwei ORFs) Minigenome verwendet (Vgl. Kapitel 3.2.3.2; Abb. 13). Hier werden sämtliche ORFs der EBOV-Proteine durch ein (oder zwei) Reporterproteine(n) ersetzt. In der Regel sind dies Renilla- oder Firefly-Luziferasen. Die essenziellen regulatorischen 3'-Leader- und 5'-Trailer-Bereiche mit Replikations- und Transkriptionspromotoren sowie Enkapsidierungssignalen bleiben unverändert erhalten. Dies gilt auch für die NP-5'-UTR, die das erste GS-Signal enthält sowie für die L-3'-UTR, die das L-GE-Signal beherbergt. Das Minigenom (MG) ist auf Expressionsplasmiden kodiert und wird in der Regel von T7-Promotoren (oder RNA-Polymerase I und II-Promotoren; Abb. 13a) kontrolliert. Am 3'-Ende befindet sich eine Ribozymsequenz des Hepatitis-Delta-Virus, die präzise und Virus-authentische 3'-Minigenomenden generiert. Im Minigenomsystem wird das MG-kodierende Plasmid in Krebszellen (u.a. HEK293-Zellen; humane embryonale Nierenzellen oder HuH7-Zellen; humane Leberzellen) transfiziert. Neben dem Minigenom werden weitere Plasmide (auch T7-Promotor-gesteuert) transfiziert, welche für die viralen Nukleokapsidproteine L, VP35, NP und VP30 kodieren sowie eine T7-RNA-Polymerase, die von einem zellulären Promotor kontrolliert wird. Letztere kann, sobald sie selbst in der transfizierten Zelle exprimiert wurde, die Überexpression des Negativstrang-MGs sowie der Nukleokapsidproteine

steuern. Somit können anschließend, analog zur Virusinfektion, Transkription und Replikation des Reporterproteins und der Minigenom-RNA stattfinden. In einer enzymatischen Reaktion wandeln die Reporterproteine bestimmte Substrate unter Freisetzung von Biolumineszenz in ihre Edukte um. Die Intensität der Biolumineszenz dient als indirektes Maß der viralen Transkription. Die Replikation und die daraus resultierenden höheren Transkriptions-Templatmengen potenzieren die Menge der gebildeten Reporterproteine (Hoenen et al., 2010). Es ist zu bedenken, dass dieses Replikations- und Transkriptions-kompetente Minigenomsystem allein die sekundäre Transkription parallel zur Replikation abbildet, da bereits zu Beginn alle viralen Proteine in ausreichender Menge vorhanden sind, bedingt durch ihre zelluläre Überexpression über Plasmide. Somit ermöglichen diese Systeme ausschließlich Aussagen über spätere Phasen des Infektionszyklus'. Allerdings entstehen keine infektiösen virusartige Partikel (iVLP), da die dafür benötigten viralen Proteine VP40, VP24 und GP nicht ebenfalls exprimiert wurden. Dies unterscheidet die klassischen Minigenomsysteme von iVLP-Systemen, die die Analyse der Primärtranskription ermöglichen (Abb. 13b, infektiöser iVLP Assay). Eine weitere Möglichkeit zur Analyse der viralen Transkription ohne Einfluss der Genom-Replikation ist die Verwendung Replikations-defizienter Minigenome. Hier macht man sich das Wissen zunutze, dass die Deletion der terminalen 55 Trailer-Nukleotide zur Elimination der *de novo*-vRNA (Genom-RNA)-Synthese führt (siehe 1.8.1.2) und somit den Replikationszyklus auf eine Runde (single-round) beschränkt. Folglich finden lediglich Antigenomsynthese und Transkription der Reporterprotein-mRNA statt, wobei die Aktivität der gebildeten Luziferase-Reporterproteine ein indirektes Maß allein der Transkription ist (Hoenen et al., 2010). Über sogenannte Virus-Rescue-Systeme (Abb. 13c) können Vollängenviren entlang des authentischen Virus-Lebenszyklus gebildet werden. Voraussetzung ist die Verwendung eines Plasmid-kodierten authentischen Vollängenantigenoms, welches für sämtliche virale Proteine kodiert. In dieser Arbeit wurden ausschließlich Replikations-/Transkriptions-kompetente (Abb. 13a) sowie Replikations-defiziente Minigenomsysteme verwendet.



**Abb. 13: EBOV-Minigenomsysteme. Farbschema wie in Abb. 5. (a)** Transkriptions- und Replikationskompetentes Minigenomsystem. Die für beide Prozesse benötigten Nukleokapsidkomponenten NP, VP35, L und VP30 sowie das ebenfalls Plasmid-kodierte Minigenom (MG; rot) werden über transfizierte Plasmide in Zellen überexprimiert (durch kotransfizierte oder ektopisch-exprimierte T7-RNA-Polymerasen oder zelluläre RNA-Polymerasen, Pol I/Pol II). Replikation und Transkription können über die enzymatische Aktivität des im MG kodierten Reporterproteins überwacht werden. **(b)** Infektiöser iVLP-Assay. Zusätzliche Supplementierung der MG-Systeme mit Plasmiden, die für GP, VP40 und VP24 kodieren, führt zur Bildung und Freisetzung von infektiösen virusartigen Partikeln (infectious virus-like particles; iVLPs), die naive Zellen infizieren können. Weil das Minigenom keine viralen Proteine kodiert, kann lediglich die Primärtranskription des Reporterproteins erfolgen (single-round-Lebenszyklus). **(c)** Filovirus-Rescuesysteme erlauben die Bildung von Volllänge-viren, indem statt eines MGs das virale Antigenom über Plasmide in Zellen eingebracht wird und so alle viralen Proteine in den Zellen exprimiert werden. Dies ermöglicht eine Neusynthese von viralen Genomen sowie den Zusammenbau und die Ausknospfung neuer Viruspartikel (Hume and Mühlberger, 2019).

## 2 FRAGESTELLUNG

Die vermehrten Ausbrüche und Epidemien in Zentral- und Westafrika in den vergangenen Jahren sowie aktuell im Kongo, unterstreichen die Notwendigkeit effektiver Therapeutika gegen das Ebola-Virus (EBOV). Auch wenn es erste präventive Vakzinierungsansätze gibt, die bereits im aktuellen Ausbruch im Kongo eingesetzt wurden, ist derzeit keine wirksame antivirale Kausaltherapie zugelassen. Wesentliche mechanistische Aspekte der Ebola-Virus-Transkription und -Replikation, d.h., die Grundlagen der Virusgenomvervielfältigung, sind bis heute unklar. Aber genau dieses Verständnis ist für die Identifikation geeigneter antiviraler Zielstrukturen unerlässlich.

Ziel der vorliegenden Studie war es, sequenzspezifische und strukturelle Voraussetzungen der Polymerisationsinitiation am 3'-Leader-Promoter sowie Mechanismen der Regulation durch das virale Protein 30 (VP30) aufzuklären. Die Besonderheit des 3'-Leader-Promotors ist, dass er sowohl die Genom-Replikation als auch die Transkription der ersten NP mRNA reguliert. Daraus lassen sich folgende zentrale Fragestellungen ableiten: (i) Gibt es überlappende Replikations- und Transkriptionspromotorsequenzen? Gibt es darüber hinaus regulatorische Sequenzbereiche, die selektiv entweder die Replikation oder die Transkription begünstigen? (ii) Welche Relevanz hat die Trennung von Replikations- und/oder Transkriptionssignalen durch den Spacer-Bereich, gibt es minimale oder maximale Längenlimits? Ist die Einhaltung der „Rule of 6“ nicht nur für die virale Replikation obligat, sondern auch für die Transkriptionsinitiation an der ersten NP-Transkriptionsstartsequenz (TSS)? (iii) Werden, ähnlich wie bei anderen NNS-Viren, Ebola-Virus-spezifische Leader-Transkripte vor Transkriptionsinitiation an der TSS synthetisiert? (iv) Welche Rolle spielt die Ausbildung der NP-Hairpin-Struktur im TSS-Spacer-Bereich für den Wechsel von der VP30-abhängigen Transkription zur VP30-unabhängigen Replikation? Gibt es strukturelle und sequenzspezifische Voraussetzungen des NP-Hairpins, die entweder die allgemeine virale RNA-Syntheserate beeinflussen oder im Speziellen die VP30-Abhängigkeit regulieren?

In der vorliegenden Arbeit wurden Replikations-/Transkriptions-kompetente und Replikations-defiziente Minigenome verwendet, die entweder monocistronisch (aus einem ORF bestehend) oder bicistronisch (zwei ORFs) waren (siehe Kapitel 3.2.3.2 und Abb. 17). Diese Minigenome ermöglichten die gezielte Mutation des Ebola-Virus-3'-Leader-Promotors, des Spacer-Bereichs, bzw. der NP-Hairpinstruktur. Mit wenigen Ausnahmen wurden die Minigenom-Assays sowohl in An- als auch in Abwesenheit des Transkriptionsfaktors VP30 durchgeführt, um Auswirkungen der eingebrachten Mutationen auf die VP30-Regulation der viralen Transkription und Replikation zu

analysieren. Rückschlüsse auf Transkriptions- und/oder Replikationsaktivität erfolgten entweder indirekt über Reporterprotein-Aktivitätsmessungen (Luziferase-Assays) oder direkt durch Quantifizierung der viralen RNA-Spezies mittels quantitativer Reverse-Transkriptase-Polymerase-Kettenreaktion, Hochdurchsatz-RNA-Sequenzierung oder Northern Blot-Analyse.

### 3 MATERIAL UND METHODEN

#### 3.1 Material

##### 3.1.1 Chemikalien

Tab. 3: Chemikalien

Name	Hersteller
2-log-ladder	NEB
2x Lysis-Juice	PJK
50 bp DNA ladder	Roth
100 bp DNA ladder	Cytogen
1 kb DNA ladder	Cytogen
Acrylamid M-Bis (24:1); 50 %	GERBU
Agar Agar	Serva
Agarose GTQ	Roth
Ammoniumpersulfat (APS)	Biorad
Ampicillin-Na (Natriumsalz)	Roth
Borsäure	Roth
Bromphenolblau	Merck
Calciumchlorid ( $\text{CaCl}_2$ )	Roth
Chloroform	Merck
Coelenterazin (CTZ)	PJK
Dinatriumhydrogenphosphat ( $\text{Na}_2\text{HPO}_4$ )	Merck
dNTP-Set (100 mM each)	Thermo Fisher Scientific
Essigsäure (100 % p.a.)	Roth
Ethanol (99,8 % p.a.)	Roth
Ethylendiamintetraessigsäure (EDTA)	GERBU
Fötales Kälberserum (FCS)	Thermo Fisher Scientific
Formamid (99,5 % p.a.)	Roth
GelRed® Nucleic Acid Gel Stain, 10.000x	Biotium

<b>Name</b>	<b>Hersteller</b>
Glycerin (99,5 %)	GERBU
GlycoBlue™ Coprecipitant	Thermo Fisher Scientific
Hefeextrakt	GERBU
Heparin	Sigma-Aldrich
Kaliumchlorid (KCl)	Merck
Kaliumdihydrogenphosphat (KH <sub>2</sub> PO <sub>4</sub> )	Roth
L-Glutamin (200 mM)	Thermo Fisher Scientific
Magnesiumchlorid	Roth
Natriumacetat	Grüssing GmbH
Natriumchlorid (NaCl)	Roth
Natriumdihydrogenphosphat (NaH <sub>2</sub> PO <sub>4</sub> )	Merck
Natriumhydroxid (NaOH)	Fluka
MOPS (3-(N-Morpholino)propanesulfonsäure)	Merck
Penicillin/Streptomycin 5000 IU/ml	Thermo Fisher Scientific
RiboLock RNase Inhibitor (40 U/μL)	Thermo Fisher Scientific
RiboRuler High Range RNA Ladder (200-6.000 bp)	Thermo Fisher Scientific
Roti®-Phenol/Chloroform/Isoamylalkohol	Roth
Rubidiumchlorid (RbCl)	Roth
Salzsäure (HCl), konz.	Roth
Spermidin	Roth
N,N,N',N'-Tetramethylethylenediamin (TEMED)	Sigma-Aldrich
TransIT® LT1 Reagent	Mirus Bio
Tris-(hydroxymethyl)-aminomethan (TRIS)	Roth
Trypton	Roth
Trypsin-EDTA (0,5 %)	Thermo Fisher Scientific
UREA	GERBU
Xylencyanol Blau	Roth



### 3.1.2 Verbrauchsmaterialien

Tab. 4: Verbrauchsmaterialien

Material	Hersteller
96-Well Platten:	
LumiNunc™, weiß	Nunc
MicroAmp™ Fast Optical 96-Well Reaction Plate und MicroAmp™ Optical Adhesive Film	Thermo Fisher Scientific
<hr/>	
CryoTube™ Vials; 1,8 mL	Thermo Fisher Scientific
Filtropur BT 50 bottle top filter, 500 mL	Sarstedt
Gilson™ DistriTips™ Spritzen	Thermo Fisher Scientific
Mikroschraubröhrchen, 1,5 mL, 2 mL	Sarstedt
Pasteurpipetten	Sarstedt
Petrischalen	Sarstedt
Pipettenspitzen mit und ohne Filter TipOne®	Sarstedt und Starlab
Reagiergefäße:	
1,5 / 2 mL	Sarstedt
0,2 mL	Biozym
Multiply®-µStrip Pro, 0.2 mL	Sarstedt
<hr/>	
Polypropylen-Reagierröhren, 15 mL, 50 mL	Greiner Bio-One
Serologische Pipetten	Sarstedt
Skalpell Cutfix Stainles #21	B.Braun
Sterican® Standardkanülen; 1 mL	B.Braun
Zellkulturflaschen (75 cm²)	Greiner bio-one
Zellkulturplatten, 6-Well	Greiner bio-one
Zellschaber	Sarstedt

### 3.1.3 Enzyme und Kits

Tab. 5: Enzyme

Enzyme	Hersteller
5x HOT CYTOpol® Blend Master Mix (7,5 mM MgCl <sub>2</sub> ; ready to use)	Cytogen
Ambion™ DNase I	Thermo Fisher Scientific
FastDigest <i>Bam</i> HI	Thermo Fisher Scientific
FastDigest <i>Bsp</i> 119I ( <i>Bst</i> BI)	Thermo Fisher Scientific
FastDigest <i>Dpn</i> I	Thermo Fisher Scientific
FastDigest <i>Eco</i> RI	Thermo Fisher Scientific
FastDigest <i>Nde</i> I	Thermo Fisher Scientific
FastDigest <i>Not</i> I	Thermo Fisher Scientific
FastDigest <i>Pac</i> I	Thermo Fisher Scientific
FastDigest <i>Xho</i> I	Thermo Fisher Scientific
Pfu DNA-Polymerase	Thermo Fisher Scientific
Phusion Master Mix 2x	Thermo Fisher Scientific
Phusion DNA-Polymerase	Thermo Fisher Scientific
PowerUp™ SYBR™ Green Master Mix 2x	Thermo Fisher Scientific
RevertAid H Minus Reverse Transcriptase	Thermo Fisher Scientific
T4 DNA Ligase	Thermo Fisher Scientific
T4 Polynukleotid-Kinase	Thermo Fisher Scientific

Tab. 6: Kits

Kits	Hersteller
Beetle-Juice BIG KIT	PJK
E.Z.N.A.® Plasmid Maxi Kit	Omega BIO-TEK
GeneJET Plasmid Miniprep Kit	Thermo Fisher Scientific
<i>mir</i> Vana™ miRNA Isolation Kit + Phenol	Thermo Fisher Scientific
NucleoBond PC 500 Purification Kit	Macherey-Nagel

<b>Kits</b>	<b>Hersteller</b>
Renilla-Juice BIG KIT	PJK
RNase-Free DNase Set	QIAGEN
RNeasy® Mini Kit	QIAGEN
Wizard® SV Gel and PCR Clean-Up system	Promega

### 3.1.4 Lösungen und Puffer

Alle für die bakterielle Zellkultur verwendeten Lösungen und Materialien wurden, sofern möglich, bei 121 °C für 20 min bei 1 bar autoklaviert. Bei Vorliegen hitze-instabiler Bestandteile wurde alternativ eine Sterilfiltration durchgeführt. Sofern nicht anders angegeben, wurde in allen Fällen ddH<sub>2</sub>O als Lösungsmittel verwendet.

Tab. 7: Lösungen

<b>Antibiotika-Lösungen</b>	<b>Zusammensetzung</b>	
Ampicillin-Lösung	100 mg/mL	Ampicillin-Na
<b>Sonstige Lösungen</b>	<b>Zusammensetzung</b>	
10% denaturierende PAA-Lösung (aus M-Bis (24:1); 50%); 1 L	200 mL	5x TBE
	200 mL	PAA
	480 g	Harnstoff (UREA)
8 M UREA-Lösung; 1 L	200 mL	5x TBE
	480 g	Harnstoff (UREA)
TBF 1-Lösung	10 mM	MOPS (pH 7)
	10 mM	RbCl
TBF 3-Lösung	100 mM	MOPS (pH 6,5)
	50 mM	CaCl <sub>2</sub>
	10 mM	KCl
	15 % (v/v)	Glycerin

Tab. 8: Puffer

Puffer	Zusammensetzung	
2x RNA-Probenpuffer (denaturierend)	0,02 % (w/v)	Bromophenolblau
	0,02 % (w/v)	Xylencyanol Blau
	8 M	Harnstoff (UREA)
	50 % (v/v)	Formamid
	2x	TBE-Puffer
5x DNA-Probenpuffer	0,25 % (w/v)	Bromophenolblau
	0,25 % (w/v)	Xylencyanol Blau
	25 % (v/v)	Glycerin
	5x	TBE-Puffer
5x TBE-Puffer	445 mM	Tris Base
	445 mM	Borsäure
	10 mM	EDTA
Phosphatpuffer defizient ( PBS <sub>def</sub> ), pH 7,5 (autoklaviert); 1 L	0,2 g	KCl
	0,2 g	KH <sub>2</sub> PO <sub>4</sub>
	1,15 g	Na <sub>2</sub> HPO <sub>4</sub>
	8 g	NaCl

### 3.1.5 Wachstumsmedien

#### 3.1.5.1 Wachstumsmedien für Bakterienzellen

Tab. 9: Wachstumsmedien für Bakterienzellen

Name	Zusammensetzung	
Lysogeny Broth (LB) Medium (autoklaviert)	10 g	Pepton aus Casein
	5 g	Hefeextrakt
	10 g	NaCl
		dH <sub>2</sub> O ad 1L
LB Agar (autoklaviert)		LB-Medium (s.o.)
	15 g	Agar

Das LB-Medium wurde mit NaOH bei Raumtemperatur (RT) auf pH 7,5 eingestellt. Vor Inokulation des Mediums mit Bakterienkolonien wurde Ampicillin in einer Konzentration von 100 mg/mL hinzugesetzt (Endkonzentration: 100 µg/mL). Die Lagerung des

Mediums erfolgte bei 4 °C. Zur Langzeitaufbewahrung von Bakterienkulturen wurde ein Volumenteil Glycerol zur Flüssigkultur hinzugefügt, gut vermischt, anschließend in flüssigem Stickstoff eingefroren und bei -80 °C gelagert. Diese Glycerol-Stocks wurden für das Ausstreichen auf Agar-Platten verwendet, von denen Einzelkolonien gepickt wurden, um frische Flüssig-LB-Medien zu inokulieren, z.B. zwecks Plasmidexpression. Der LB-Agar wurde für das Ausplattieren und die Kultur von Bakterienzellkulturen in Petrischalen verwendet. Dafür wurde der LB-Agar zunächst in der Mikrowelle erhitzt und verflüssigt. Nach Abkühlung auf unter 55 °C wurde das Medium mit Ampicillin in einer Endkonzentration von 100 µg/mL versetzt und auf Petrischalen verteilt.

### 3.1.5.2 Wachstumsmedien für Säugerzellen

Sämtliche Wachstumsmedien wurden von der Firma Thermo Fisher Scientific (TFS) bezogen.

Tab. 10: Wachstumsmedien für Säugerzellen

Name	Zusammensetzung	
OPTI-MEM®	Siehe Herstellerangaben	
Dulbecco's Modified Eagle Medium (DMEM)	Siehe Herstellerangaben	
DMEM <sub>10%FCS+Q+P/S</sub>	500 mL	DMEM
	50 mL	FCS (10 %)
	5 mL	L-Glutamin (2 mM)
	5 mL	Penicillin/Streptomycin (50 IU/mL; 50 µg/mL)

### 3.1.6 Zellen

#### 3.1.6.1 Prokaryotische Zellen

Tab. 11: Prokaryotische Zellen

Bakterienstamm	Quelle	Identifikationsnr.
<i>E. coli</i> DH5α	N/A	N/A

### 3.1.6.2 Eukaryotische Zellen

Tab. 12: Eukaryotische Zellen

Zelllinie	Quelle	Identifikationsnr.
Humane embryonale Nierenzellen (HEK293)	ATCC	DMSZ ACC 305
humane Hepatomazelllinie (HuH7)	ATCC	JCRB 0403

### 3.1.7 Viren

Tab. 13: Viren

Name	Quelle	Identifikationsnr.
EBOV; Stamm: Mayinga; 1976	Institut für Virologie, Marburg	AF086833

### 3.1.8 Nukleinsäuren

#### 3.1.8.1 DNA-Oligonukleotide (Primer)

Alle aufgeführten Primer wurden von der Firma Metabion synthetisiert. Lediglich das Random-Hexamer-Primer-Set wurde von TFS bezogen. Alle für diese Arbeit verwendeten spezifischen Klonierungsprimer sind den jeweiligen Manuskripten sowie den Ergänzungen (Supplementaries) zu entnehmen (Kapitel 4).

Tab. 14: RT-Primer

Name	Sequenz (5' → 3')	Amplifikation von
Random Hexamer Primers	N/A	Firefly-Luciferase mRNA
pS_0133_luc(+)	GGC CTC TTC TTA TTT ATG GCG A	MG vRNA
pS_0134_luc(-)	AGA ACC ATT ACC AGA TTT GCC TGA	MG mRNA+cRNA
pS_0140_RT_cRNA	CAG TCC TGC CTT TTC TTT TAA TTT TAT C	MG cRNA (Bindung an cTrailer-Region)
RT_leader	ATC TTC CTC ATA GTT ATT CGC ACA CA	MG und EBOV leaderRNA+cRNA
RT_cRNA (NP-5'-UTR)	CAA ACC AGG TGT GAT TAC AGT AAC AAT T	MG und EBOV cRNA (Bindung an NP-5'-UTR)

Tab. 15: qPCR-Primersets

Name	Sequenz (5' → 3')	Amplifikation von
pS_0133_luc(+)	GGC CTC TTC TTA TTT ATG GCG A	MG vRNA und MG mRNA+cRNA
pS_0134_luc(-)	AGA ACC ATT ACC AGA TTT GCC TGA	
pS_0140_RT_cRNA s3	CAG TCC TGC CTT TTC TTT TAA TTT TAT C	MG cRNA (cTrailer)
pS_0139_qPCR_cRNA s3	CGG TGA TAG CCT TAA TCT TTG TG	
pS_0147_qPCR_FF_fwd	CGT GCA AAA GAA GCT ACC G	Firefly-Luciferase mRNA
pS_0148_qPCR_FF_rev	GGT GGC AAA TGG GAA GTC AC	
qPCR_leader+cRNA	GCC GGA CAC ACA AAA AGA AAG AA	MG und EBOV leaderRNA+cRNA
RT_leader+cRNA	ATC TTC CTC ATA GTT ATT CGC ACA CA	
qPCR_NP mRNA	GTA ACT CAA TAT TCT AAC TAG CGA TTT ATC TAA ATT AAA TTA CA	EBOV mRNA
RT_NP mRNA	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTC TTA ATT ATA AAA CG	
qPCR_leader+cRNA_fwd	GCC GGA CAC ACA AAA AGA AAG AA	EBOV cRNA (NP-5'-UTR)
RT_cRNA s1/s2	CAA ACC AGG TGT GAT TAC AGT AAC AAT T	

Tab. 16: Sequenzierprimer

Name	Sequenz (5' → 3')	Verwendungszweck
pS_0001	GCT ACT GGC TCA ATA TGT GGC	pA 3E5E mono-Derivate
3_Seq	ACC GCC TTT GAG TGA GCT GA	pA 3E5E mono-/bici-Derivate
4_Seq	CAA TAA CAA TAT GAG CCC AGA CC	pA 3E5E mono-Derivate
5_Seq	CCT AAC AAT TTG TAC AAC GTC AGG	pA 3E5E mono-/bici-Derivate
pS_0085	GCT GAA AGG AGG AAC TAT ATC CG	pA 3E5E mono-/bici-Derivate

Name	Sequenz (5' → 3')	Verwendungszweck
pS_0086	TGA TGC CGC ATA GTT AAG CCA GC	pA 3E5E mono-/bici-Derivate
pS_0087	AGG ATC TAG GTG AAG ATC CT	pA 3E5E mono-/bici-Derivate
pS_0050	CGT AGC GCT TCA TGG CTT TG	pA 3E5E bici-Derivate
pS_0076	GCA GGG AGA GAG GCT AAA TAT AG	pA 3E5E bici-Derivate
pS_0082	GTT GTT AAC GTA GCC GCT CAT G	pA 3E5E bici-Derivate
pS_0084	TGG TAT GAG ACA TAC TCG GAA T	pA 3E5E bici-Derivate

Tab. 17: Kolonie-PCR-Primer

Name	Sequenz (5' → 3')	Verwendungszweck
1_fwd	CTT TCG AAG TCA TGG TGG TAT GAG AC	pA 3E5E mono-Derivate
2_rev	GTT AGC AGC CGG ATC CTT TTT TTT G	pA 3E5E mono-Derivate
17_*pA3E5E	TTT CTC GAG CTT GTT AGA CCA GCT TTT CTT TTA CAT TG	pA 3E5E bici-Derivate
15_pCAGGS	TCG CGG CCG CAT GAG C	pA 3E5E bici-Derivate



### 3.1.8.2 RNA-Oligonukleotide

Tab. 18: RNA-Oligonukleotide

Name	Sequenz (5' → 3')	Quelle
5'-PPP- <i>leader</i> RNA (65-mer)	pppCGG ACA CAC AAA AAG AAA GAA GAA UUU UUA GGA UCU UUU GUG UGC GAA UAA CUA UGA GGA AGA UU	AXOlabs
(+) RNA 56-158 ( <i>in vitro</i> -Transkript)	GAG GAA GAU UAA UAA UUU UCC UCU CAU UGA AAU UUA UAU CGG AAU UUA AAU UGA AAU UGU UAC UGU AAU CAC ACC UGG UUU GUU UCA GAG CCA CAU CAC AAA G	(Schlereth et al., 2016)
(+) RNA 2-78 ( <i>in vitro</i> -Transkript)	GGA CAC ACA AAA AGA AAG AAG AAU UUU UAG GAU CUU UUG UGU GCG AAU AAC UAU GAG GAA GAU UAA UAA UUU UCC UCU	(Schlereth et al., 2016)
(-) RNA 1-154 ( <i>in vitro</i> -Transkript)	GTG ATG TGG CTC TGA AAC AAA CCA GGT GTG ATT ACA GTA ACA ATT TCA ATT TAA ATT CCG ATA TAA ATT TCA ATG AGA GGA AAA TTA TTA ATC TTC CTC ATA GTT ATT CGC ACA CAA AAG ATC CTA AAA ATT CTT CTT TCT TTT TGT GTG TCC G	(Schlereth et al., 2016)

### 3.1.9 Vektoren und rekombinante Plasmide

Die nachfolgenden rekombinanten Plasmide 4 - 8 wurden im Rahmen des EBOV-spezifischen Minigenomassays (3.2.3.2) transfiziert und dienten der Koexpression der Transkriptions- und Replikations-relevanten Nukleokapsidkomponenten L, VP35, NP und ggf. VP30 sowie einer T7-Polymerase, welche das erste (mini)genomische (-) RNA-Templat generiert. Die transfizierten Minigenom-Plasmide stehen unter der Kontrolle eines T7-Promotors und stellen entweder rekombinante Derivate des monocistronischen pANDY 3E5E- (Abb. 15; Tab. 19) oder des pANDY 3E5E dtrl-Plasmids dar. Im letztgenannten Plasmid sind die terminalen 55 5'-Trailer-Nukleotide deletiert (dtrl). Im Falle der bicistronischen Minigenomsysteme wurden Derivate eines Wildtyp (Wt)-bicistronischen Minigenoms (pA 3E5E\_Rluc\_Fluc; Abb. 15) verwendet, welches aus den Plasmiden 10 und 11 sowie dem Wt-monocistronischen Minigenom pANDY 3E5E (Nr. 1) generiert wurde. Das pGL4.13-Plasmid (Nr. 9; unter der Kontrolle eines SV40-Promotors) kodiert für eine Firefly-Luziferase (luc2/Fluc), welche in monocistronischen Minigenomsystemen zur Normalisierung eingesetzt wurde. Die genauen Klonierungsstrategien sind den jeweiligen Material- und Methodenkapiteln sowie den Ergänzungen (Supplementaries) der unter Kapitel 4 aufgeführten Manuskripte zu entnehmen.

Tab. 19: Rekombinante Plasmide und Vektoren

Nr.	Plasmidname	Quelle	Identifikationsnr.
1	pCAGGS	(Hoenen et al., 2006)	N/A
2	pANDY 3E5E	(Hoenen et al., 2006)	N/A
3	pANDY 3E5E dtrl	(Hoenen et al., 2010)	N/A
4	pCAGGS-VP30	(Hoenen et al., 2006)	N/A
5	pCAGGS-NP	(Hoenen et al., 2006)	N/A
6	pCAGGS-VP35	(Hoenen et al., 2006)	N/A
7	pCAGGS-L	(Hoenen et al., 2006)	N/A
8	pCAGGS-T7	(Hoenen et al., 2006)	N/A
9	pGL4.13 [luc2/SV40]	Promega	# E6681
10	pCAGGS-luc2	(Biedenkopf and Hoenen, 2017)	N/A
11	pA-3E5E-GLuc-GFP	N. Biedenkopf; Institut für Virologie, Marburg; Siehe Kapitel 8.2	N/A

**3.1.10 Geräte****Tab. 20: Geräte**

<b>Instrumente</b>	<b>Hersteller</b>
Autoklav V 95	Systec
Destamat	Heraeus
Elektrophorese Power Supply PowerPac 3000	Biorad
Geldokumentationssysteme:	
Dark Hood DH-50 und Gerix 1000	Biostep
ChemiDoc™ MP Imaging System	Biorad
Gelkammern:	
Mini-Sub® Cell GT Systems	Biorad
PAA-Gelkammern, Spezialanfertigung	Universität Lübeck
Heizblöcke:	
Thermoblock TB1	Biometra
Neoblock 1	neoLab®
Inkubatoren:	
Brutschrank BE 400	Memmert
Schüttelinkubator 3033	GFL
HERAcell150	Thermofisher Scientific
HERAcell240	Thermofisher Scientific
<hr/>	
Lichtmikroskop Wilovert®	Will
Luminometer Centro LB 960	Berthold
Magnetrührer	Heidolph Instruments
Mikroplattenluminometer Centro LB 960	Berthold
Mikrowelle 8020 E	Privileg
Gilson™ Pipettiersystem	Thermo Fisher Scientific
PCR-Cycler:	
TGradient	Biometra
Tone 96 G	Biometra
<hr/>	
pH-Meter inoLab pH Level 1	WTW
Pipetten	Eppendorf

<b>Instrumente</b>	<b>Hersteller</b>
Pipettierhilfe pipetus®-akku	Integra Bioscience
QuantStudio3™ Real-Time PCR System	Thermo Fisher Scientific
Sicherheitswerkbank Heraeus Herasafe®	Thermo Fisher Scientific
Spektralphotometer:	
UV5Nano	Mettler Toledo
Nanodrop ND1000	Peqlab
<hr/>	
Thermomixer 5436	Eppendorf
Vortex Schüttler 7-2020	neoLab
Waagen	Sartorius
Wasserbad MT	Lauda
Zentrifugen:	
Minispin® plus	Eppendorf
Microcentrifuge 5415D	Eppendorf
Centrifuge 5415R	Eppendorf
Centrifuge 5810R	Eppendorf
Biofuge Fresco	Heraeus

### 3.1.11 Verwendete Software

Tab. 21: Verwendete Software

<b>Name der Software</b>	<b>Verwendungszweck</b>
Adobe Illustrator CS6	Bildbearbeitung
Affinity Designer 1.7.2.	Bildbearbeitung
GraphPad Prism version 8.1.1	Statistische Auswertung, Datendarstellung
Inkscape	Bildbearbeitung
Microsoft Office 2016	Auswertung und Texterstellung
RNAfold (Vienna Web-Package)	Faltung von RNA-Sekundärstrukturen
Snappgene	Klonierungsmanagementsoftware

## 3.2 Methoden

### 3.2.1 Molekularbiologische Methoden

#### 3.2.1.1 Amplifikation von DNA durch Polymerase-Kettenreaktion (PCR)

Standard-PCRs wurden zur Amplifikation und *in vitro*-Mutagenese von DNA im Rahmen von Klonierungen durchgeführt. Als DNA-Polymerasen wurden entweder die Pfu-DNA-Polymerase oder die Phusion-DNA-Polymerase (beide TFS) eingesetzt. Pufferbedingen, Annealingtemperaturen, Elongationszeiten usw. wurden entsprechend der Herstellerangaben angepasst. Allgemein wurden 30 PCR-Amplifikationszyklen nach unten aufgeführtem Schema durchgeführt. Im Anschluss an die PCR wurden die resultierenden Fragmente per Agarosegelelektrophorese (3.2.1.12) überprüft und anschließend über das Wizard®SV Gel and PCR Clean-Up Systems (3.2.1.11) eluiert und aufgereinigt.

Tab. 22: Standard PCR

Bestandteile	Endkonzentration
DNA-Templat	0,6 ng/μL
10x Pfu-Puffer / 5x HF Puffer	1x
dNTP-Mix	0,4 mM (pro dNTP)
Primer forward (fwd)	1 pmol/μL
Primer reverse (rev)	1 pmol/μL
DNA-Polymerase	x U (wie für Enzym empfohlen)
ddH <sub>2</sub> O	ad 50 μL

Tab. 23: Standard PCR-Programm

Schritt	Zeit	Temperatur
Initiale Denaturierung	2 min	98 °C
Denaturierung	10 - 30 s	98 °C
Primerannealing	30 s	b
Elongation	a	72 °C
Finale Elongation	7 min	72 °C

a) Abhängig von Produktlänge und Elongationsrate der verwendeten Polymerase;

b) Abhängig von der Schmelztemperatur der verwendeten Primer und der Polymerase.

### 3.2.1.2 Kolonie-PCR

Um nach Klonierungen zu überprüfen, ob das gewünschte Insert in der Plasmid-DNA enthalten ist, wurde eine Kolonie-PCR durchgeführt. Hierfür wurde eine Bakterienkolonie von der LB-Agar-Platte mit einer Pipettenspitze gepickt und je nach Größe in 30 – 50 µL ddH<sub>2</sub>O resuspendiert. Die Suspension wurde anschließend bei 98 °C für 5 min erhitzt. Ein µL davon wurde nachfolgend in der PCR-Reaktion eingesetzt. In allen Fällen wurde der 5x HOT CYTOpol® Blend Master Mix (7,5 mM MgCl<sub>2</sub>; ready to use) von Cytogen nach Herstellerangaben verwendet und eine Amplifikation über 25 Zyklen durchgeführt. Der HOT CYTOpol® Mix ermöglicht das direkte Auftragen der PCR-Produkte auf ein analytisches Agarosegel (3.2.1.12) im Anschluss an die PCR. Die verwendeten Kolonie-PCR-Primer für monocistronische und bicistronische Minigenomplasmide sind unter (3.1.8.1) aufgeführt.

Tab. 24: Kolonie-PCR-Ansatz

Komponente	Endkonzentration
Kolonie (in 50 µL ddH <sub>2</sub> O)	/
Primer fwd (10 µM)	1 pmol/µL
Primer rev (10 µM)	1 pmol/µL
5x HOT CYTOpol® Master Mix	1x
ddH <sub>2</sub> O	Ad 20 µL

Tab. 25: Kolonie-PCR-Programm

Schritt	Zeit	Temperatur
Initiale Denaturierung	5 min	98 °C
Denaturierung	30 s	98 °C
Primerannealing	10 s	b
Elongation	a	72 °C
Finale Elongation	10 min	72 °C

a) Abhängig von Produktlänge und Elongationsrate der verwendeten Polymerase;

b) Abhängig von der Schmelztemperatur der verwendeten Primer und der Polymerase.

### 3.2.1.3 Quantitative Reverse-Transkriptase-Polymerase-Kettenreaktion (qRT-PCR)

Zwei-Schritt-qRT-PCR-Analysen wurden durchgeführt, um die Menge viraler RNA-Spezies (vRNA, cRNA, mRNA und ggf. *leader*RNA) nach Infektion mit EBOV (3.1.7 und 3.2.3.1) oder nach Transfektion des Minigenomsystems (3.2.2.2; 3.2.3.2) in der zellulären (HuH7- oder HEK293-Zellen; 3.1.6.2) Gesamt-RNA zu quantifizieren. Im ersten Schritt der qRT-PCR, der Reversen Transkription, wurden die RNAs zunächst anhand des nachfolgenden Protokolls (Tab. 26; Tab. 27) mithilfe der RevertAid H Minus Reversen Transkriptase (TFS) in cDNA umgeschrieben. Die entsprechenden RT-Primer (3.1.8.1) binden antisense der zu detektierenden RNA-Spezies. Die Bindestellen sowie spezifische qRT-PCR-Strategien sind in den Einzelmanuskripten (Kapitel 4) oder ihren Ergänzungen (Supplementaries) aufgeführt. Für jede RNA-Probe wurde eine Reaktion mit (+RT) und ohne (-RT) Reverse Transkriptase durchgeführt, um unspezifische Amplifikationen (z.B. durch DNA-Kontaminationen) zu detektieren.

Tab. 26: Pipettierschema der Reversen Transkription.

Komponente	Reaktion +RT	Reaktion -RT
Endkonzentration		
Gesamt-RNA	25 ng/μL	
RT-Primer (10 μM)	0,5 pmol/μL	
ddH <sub>2</sub> O	Ad 12,2 μL	Ad 13,2 μL
Inkubation bei 65 °C für 5 min. Danach Abkühlen auf 25 °C.		
dNTPs (je 25 mM)	1 nmol/μL	
5x RT-Puffer	1x	
RevertAid H minus RT (200 U/μL)	200 U	/
20 μL		

Tab. 27: Programm der Reversen Transkription.

Zeit	Temperatur
10 min	25 °C
60 min	42 °C
10 min	70 °C

Im Anschluss an die Reverse Transkription erfolgte die Quantifizierung der cDNAs (1:10 verdünnt) per quantitativer PCR mithilfe des QuantStudio3™ Real-Time PCR Systems unter Verwendung des PowerUp™ SYBR™ Green Master Mixes (MM; beides TFS) im Fast-Cycling-Modus (Tab. 29) nach Herstellerangaben. Die qRT-PCRs mit Gesamt-RNA aus EBOV-infizierten HuH-7-Zellen wurden von Dr. Julia Schlereth (Institut für Pharmazeutische Chemie, Marburg) durchgeführt.

Tab. 28: qPCR-Pipettierschema.

Komponente	Endkonzentration
cDNA	~ 5 ng/μL
Primer fwd (10 μM)	1 pmol/μL
Primer rev (10 μM)	1 pmol/μL
2x PowerUp™ SYBR™ Green MM	1x
ddH <sub>2</sub> O	Ad 20 μL

Tab. 29: qPCR Fast-Cycling-Programm.

Schritt	Zeit	Temperatur	
UDG-Aktivierung	2 min	50 °C	
Polymeraseaktivierung	2 min	98 °C	
Denaturierung	1 s	98 °C	40 Zyklen
Annealing/Elongation	30 s	60 °C	

Die relative Quantifizierung erfolgte unter Verwendung des DNA-Farbstoffinterkalators SYBR Green. Wenn der Farbstoff in die DNA interkaliert, verändert er sich strukturell und wird weniger mobil. Die überschüssige Energie, die dabei entsteht, wird als Fluoreszenz abgegeben. Diese steigt mit zunehmender Menge an amplifizierter DNA. Die qPCR wurde in 40 Zyklen durchgeführt, wobei nach jedem Zyklus die Fluoreszenzintensität detektiert wurde. Rückschlüsse auf die jeweiligen RNA-Mengen erfolgten mittels des  $2^{-\Delta\Delta C_T}$ -Algorithmus.  $C_T$ -Werte (= cycle threshold) geben die Zykluszahl an, bei dem die Fluoreszenzintensität das erste Mal über dem Schwellenwert (Quantifikationslimit) lag. Genaue Vorgehensweisen der  $2^{-\Delta\Delta C_T}$ -Quantifizierung sind den Manuskripten in Kapitel 4 zu entnehmen.



### 3.2.1.4 Klonierungsstrategien

In dieser Studie wurden alle rekombinanten Plasmide in das Rückgrat des Replikations-/Transkriptions-kompetenten (RC) monocistronischen Minigenoms pANDY 3E5E (Abb. 15, oben) oder in das Replikations-defiziente (RD) Derivat pANDY 3E5E dtrl (Deletion der 55 5'-terminalen Trailer-Nukleotide) kloniert. Auch das bicistronische Minigenom pA 3E5E\_Rluc\_Fluc (abgekürzt pA 3E5E\_bici) basiert auf dem pANDY 3E5E-Plasmid (Abb. 15, unten). Das (-) RNA-Genom der mono- oder bicistronischen Minigenom-Derivate ist schematisch unter Punkt (3.2.3.2) dargestellt. Es wurden entweder klassische PCR-Restriktionsklonierungen (siehe unten) über Restriktionsenzyme (alle relevanten in Tab. 5 angegeben) oder eine ortsgerichtete *in vitro*-Mutagenesen (siehe unten) durchgeführt. Alle spezifischen Klonierungsstrategien sowie spezifische Klonierungsprimer sind den Material- und Methodenkapiteln sowie den Ergänzungen (Supplementaries) der Einzelmanuskripte (Kapitel 4) zu entnehmen.

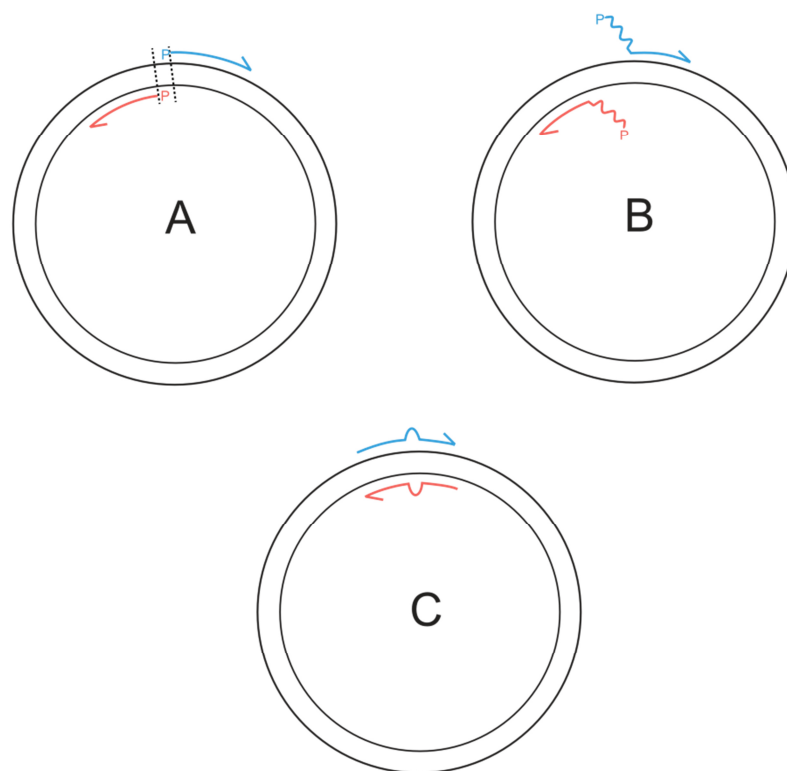
#### Restriktionsklonierung

Im Rahmen von Restriktionsklonierungen wurde das gewünschte Insert inklusive der relevanten Erkennungssequenz von Restriktionsenzymen mittels PCR zunächst amplifiziert und ggf. aufgereinigt (3.2.1.11). Anschließend wurde sowohl das PCR-Insert als auch der Zielvektor mit den gleichen Restriktionsenzymen geschnitten (3.2.1.5), und dabei zueinander passende 3'- und 5'-Überhänge generiert, die eine nachfolgende Ligation (3.2.1.8) des Inserts mit dem Rückgrat des Zielvektors ermöglichten (Zirkularisierung). Nach der Ligation wurde das resultierende Plasmid in chemisch-kompetente *E. coli* DH5 $\alpha$ -Zellen transformiert (3.2.1.10) und anschließend auf Ampicillin-LB-Agarplatten selektiert (Plasmide tragen Amp-Resistenzkassette).

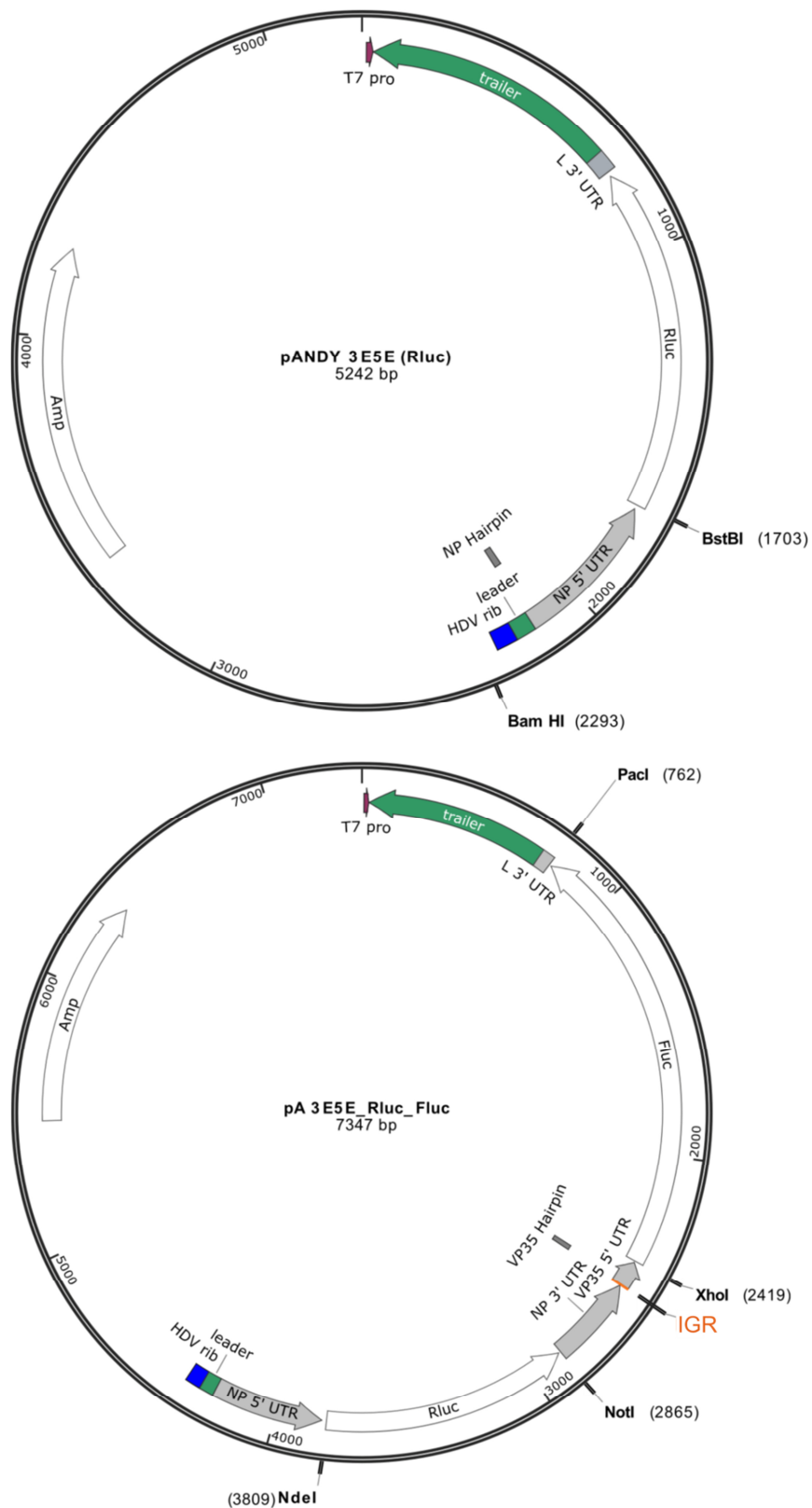
#### Ortsgerichtete *in vitro*-Mutagenese

Ortsgerichtete *in vitro*-Mutagenese wurde genutzt, um Nukleotidsequenzen in Plasmid-kodierte Minigenome (Abb. 15) über spezifische Primer entweder zu deletieren, zu mutieren oder um bestimmte Sequenzen zu insertieren. Im Wesentlichen fanden dabei die nachfolgenden drei Strategien (A - C) Anwendung (Abb. 14). Strategie A wurde genutzt, um Deletionen durchzuführen. Strategie B diente der Insertion von Nukleotiden durch Überhang-Primer. Strategie C nutzt zueinander komplementäre Primer, die kleine Mismatches (1 – 5 Nt) erlauben, über die entweder Insertionen oder Deletionen in die Plasmid-DNA eingefügt wurden. Im Anschluss an die jeweiligen PCR-Reaktionen und ihre Aufreinigung (3.2.1.11) wurden die PCR-Produkte einem *DpnI*-Verdau (3.2.1.6) unterzogen, um das Vektor-Templat zu eliminieren. In den Fällen A und B wurden die

PCR-Produkte im Anschluss phosphoryliert (3.2.1.7) und daraufhin ligiert, bevor sie in chemisch-kompetente *E. coli* DH5 $\alpha$ -Zellen transformiert (3.2.1.10) und auf Ampicillin-LB-Agarplatten selektiert wurden. Die Strategie C erforderte keine Phosphorylierung der PCR-Produkte. Hier fand die Ligation direkt nach der Transformation in den Bakterienzellen selbst statt. Positive Klone (Amp-Selektion) wurden in beiden Klonierungsstrategien verwendet, um 3 – 6 mL LB-Flüssigkulturen (3.2.2.1) für die Plasmidexpression anzupflanzen. Vorab wurde eine Kolonie-PCR (3.2.1.2) durchgeführt, um das Vorhandensein, bzw. die Größe des Inserts, zu überprüfen. Aufgereinigte Plasmide (3.2.1.11) wurden per DNA-Sequenzierung (3.2.1.15) auf die Richtigkeit der Insertsequenz überprüft.



**Abb. 14: Strategien der ortsgerichteten *in vitro*-Mutagenese in Kombination mit *DpnI*-Restriktionsverdau.** (A) Deletionsmutagenese über Rücken-an-Rücken-Primer. (B) Insertions-Mutagenese über Rücken-an-Rücken-Überhangprimer. (C) Komplementäre-Primer-Mutagenese zur Insertion oder Deletion von Nukleotiden. Hier dargestellt als Insertion. Strategien A und B erfordern eine Primer- oder PCR-Produkt-Phosphorylierung sowie eine *in vitro*-Ligation. In Strategie 2 ist keine Phosphorylierung vonnöten. Die Ligation wird in der Bakterienzelle durchgeführt.



**Abb. 15: Minigenom-kodierende Plasmide** (monocistronisch und bicistronisch). Die Sequenz der Minigenome (monocistronisch: pANDY 3E5E; bicistronisch: pANDY 3E5E\_Rluc\_Fluc) ist schematisch in Abb. 17 dargestellt und detailliert unter 3.2.3.2 erläutert.

### 3.2.1.5 Restriktionsverdau von DNA

Im Rahmen von Restriktions-Klonierungen rekombinanter Plasmide wurden doppelsträngige (ds) PCR-DNA-Fragmente oder Vektor-DNAs endonukleolytisch gespalten. Die verwendeten Restriktionsendonukleasen erkennen spezifische palindromische Sequenzen und spalten an entsprechender Stelle die Phosphodiesterbindungen der DNA hydrolytisch auf. Die verwendeten Endonukleasen sind dem Kapitel Material Enzyme und Kits (3.1.3) zu entnehmen. Der Verdau wurde nach jeweiliger Empfehlung des Herstellers durchgeführt. Die Aufreinigung von Plasmid-Vektoren nach Restriktionsverdau erfolgte zunächst über ein präparatives Agarosegel mit anschließender Gelelution und Aufreinigung über das Wizard®SV Gel and PCR Clean-Up System (Promega) und im Falle eines PCR-Produkt-Restriktionsverdaus direkt mithilfe des Wizard®SV Gel and PCR Clean-Up Systems (3.2.1.11). PCR-Produkte wurden in 20 µL ddH<sub>2</sub>O und Plasmid-DNA in 50 µL eluiert.

### 3.2.1.6 DpnI-Verdau von Templat-Vektor

Im Rahmen der ortsgerichteten *in vitro*-Mutagenese (3.2.1.4) wurde ein *DpnI*-Verdau zur Entfernung der Templat-Vektor-DNA aus dem PCR-Mutageneseansatz durchgeführt. Der Reaktionsansatz wurde für 30 – 60 min bei 37 °C inkubiert. Sofern für das weitere Vorgehen erforderlich, wurde im Anschluss eine Hitzeinaktivierung für 5 Minuten bei 80 °C durchgeführt. Der *DpnI*-Verdau wurde per 1 %-Agarosegelelektrophorese (3.2.1.12) überprüft.

Tab. 30: Ansatz *DpnI*-Restriktionsverdau

Komponente	Endkonzentration/Volumen
Insert-PCR-Produkt	~ 10 ng/µL
10x FastDigest Puffer	1x
FastDigest- <i>DpnI</i>	0,5 µL
ddH <sub>2</sub> O	Ad 20 µL

### 3.2.1.7 Phosphorylierung von DNA

PCR-Produkte wurden im Rahmen von ortsgerichteten *in vitro*-Mutagenesen (3.2.1.4) nach *DpnI*-Verdau (3.2.1.6) und vor ihrer Ligation (3.2.1.8) mithilfe der T4-Polynukleotidkinase (TFS) phosphoryliert. Der nachfolgende Reaktionsansatz wurde für 20 min bei 37 °C inkubiert und anschließend bei 75 °C für 10 min inaktiviert.

Tab. 31: Ansatz DNA-Phosphorylierung

Komponente	Endkonzentration
<i>DpnI</i> -verdautes Insert	~ 12,5 ng/μL
10x buffer A (T4 PNK)	1x
dATP (10 mM)	1 mM
T4 PNK (10 U/μL)	10 U
ddH <sub>2</sub> O	Ad 20 μL

### 3.2.1.8 Ligation von DNA-Fragmenten

Phosphorylierte PCR-Produkte (3.2.1.7) sowie restriktionsverdaute Plasmid-DNA oder restriktionsverdaute PCR-Produkte mit 3'- und 5'-Überhängen (3.2.1.5) wurden im Rahmen von ortsgerichteten *in vitro*-Mutagenesen oder Standard-Restriktions-Klonierungen ligiert. Während der Ligation entstehen neue Phosphodiesterbrücken zwischen den dsDNA-Enden. Die Ligation mittels T4 DNA-Ligase (TFS) wurde bei Raumtemperatur für 50 min durchgeführt.

Tab. 32: Ligations-Ansatz

Komponente	Enkonzentration
verdaute, gereinigte Insert-DNA	~5 ng/μL
verdaute, gereinigte Vektor-DNA	~2 ng/μL
10x T4 Ligase-Puffer	1x
T4 DNA-Ligase (5U/μL)	1 U
ddH <sub>2</sub> O	Ad 20 μL

### 3.2.1.9 Präparation chemisch-kompetenter *E.coli* DH5α-Zellen

Chemisch-kompetente *E. coli* DH5α-Zellen wurden im Rahmen von Klonierungen (3.2.1.4) eingesetzt. Sie sind in der Lage rekombinante Plasmide über die sogenannte *Transformation* (3.2.1.10) aufzunehmen. *E. coli*-Zellen wurden zunächst in einer 3 mL LB-Flüssigkultur (Vorkultur) über Nacht bei 37 °C und 180 - 200 rpm schüttelnd inkubiert. Die Vorkultur diente am Folgetag zum Animpfen der Hauptkultur von 150 mL LB-Medium. Diese wurde unter gleichen Wachstumsbedingungen bis zu einer OD<sub>600</sub> von 0,6 inkubiert. Die Zellen wurden per Zentrifugation bei 4 °C und 5000 rpm (7 min)

pelletiert und anschließend in 15 mL einer gekühlten TFB 1-Lösung (3.1.4) resuspendiert. Diese Suspension wurde erneut wie zuvor zentrifugiert. Die Resuspendierung des Zellpellets erfolgte in TFB 3-Lösung (3.1.4). Zuletzt wurden die chemisch-kompetenten Zellen aliquotiert, in flüssigem Stickstoff schockgefroren und bis zu ihrer Verwendung bei -80 °C gelagert.

#### **3.2.1.10 Transformation von Plasmid-DNA in Bakterienzellen**

Minigenom-kodierende Plasmide (3.1.9; Abb. 15) wurden in chemisch-kompetente *E. coli* DH5 $\alpha$ -Zellen (3.2.1.9) transformiert. Dafür wurden die *E. coli* DH5 $\alpha$ -Zellen zunächst für einige Minuten auf Eis aufgetaut. Anschließend wurden 4 – 10  $\mu$ L der zu transformierenden Plasmidlösung mit 50  $\mu$ L der chemisch-kompetenten Zellen vorsichtig vermischt und für 5 – 10 Minuten auf Eis inkubiert. Die Aufnahme der Plasmid-DNA erfolgte durch Hitzeschock bei 42 °C für 90 s. Danach wurde der Ansatz erneut für 1 min auf Eis gelagert, bevor 700  $\mu$ L Antibiotika-freies LB-Flüssignährmedium (3.1.5.1) hinzugesetzt wurden. Die im Nährmedium aufgenommenen Zellen wurden nun für 30 – 50 min bei 37 °C leicht schüttelnd inkubiert und anschließend bei 5.000 rpm für 2 min pelletiert. Das Pellet wurde in 100  $\mu$ L frischem LB-Medium resuspendiert, auf LB-Agarplatten (mit 100  $\mu$ g/ml Ampicillin als Selektionsmarker) ausplattiert und über Nacht bei 37 °C inkubiert. Alle verwendeten Minigenom-kodierenden Plasmide weisen eine Ampicillin (Amp)-Resistenzkassette auf.

#### **3.2.1.11 Isolierung und Aufreinigung von Nukleinsäuren**

##### **Isolierung und Aufreinigung von Plasmid-DNA aus Bakterienzellen**

Rekombinante Plasmide wurden mithilfe des GeneJet Plasmid Miniprep Kits (TFS), des NucleoBond PC 500 Purification Kits (Macherey-Nagel) oder des E.Z.N.A.® Plasmid Maxi Kits (Omega BIO-TEK) aus transformierten Bakterienzellen isoliert und aufgereinigt. Das Isolierungsprinzip beruht auf der Bindung von Plasmid-DNA an eine Silica-Säulenmembran. Die Anzucht von Bakterienzellen für Minipreps oder Vorkulturen für Maxipreps erfolgte in 3 oder 6 mL Flüssig-LB-Medium (inkl. 0,1 % Ampicillin; 3.1.5.1), welches mit einer Bakterienkolonie (von Amp-Platte oder aus Glycerin-Stock) angeimpft und für 6 – 8 h bei 37 °C schüttelnd (180 – 200 rpm) inkubiert wurde. Für Maxipreps wurden 200 – 300 mL LB-Medium (inkl. 0,1 % Ampicillin) mit einer Vorkultur à 3 mL angeimpft und 12 – 16 h (über Nacht) analog inkubiert. Die im Kit enthaltenen Puffer, Säulen und Filter wurden angewendet wie vom Hersteller empfohlen. In allen Fällen wurde die Plasmid-DNA in einem passenden Volumen ddH<sub>2</sub>O eluiert.

## Aufreinigung von DNA

DNA-PCR-Produkte (3.2.1.1) oder restriktions-verdaute Plasmid-DNA (3.2.1.5) wurden stets über präparative Agarosegele aufgetrennt (3.2.1.12) und anschließend mit Hilfe des Wizard SV Gel and PCR Clean-Up systems (Promega) nach Herstelleranleitung aus dem Gel eluiert, aufgereinigt und in 22 µL (PCR-Produkte; Säulentotvolumen: 2 µL) oder 50 µL (Plasmid-DNA) ddH<sub>2</sub>O eluiert. Die Aufreinigung erfolgt hier über eine Säulenmatrix. Restriktions-verdaute PCR-Produkte wurden ebenfalls mit dem Wizard SV Gel and PCR Clean-Up system aufgereinigt, allerdings ohne erneut auf einem präparativen Agarosegel aufgetragen zu werden.

## Isolierung und Aufreinigung von RNA

### *Aufreinigung über eine Silikamembran*

Die Aufreinigung zellulärer Gesamt-RNA und RNA <200 Nt aus EBOV-infizierten HuH7-Zellen (3.2.3.1) oder Minigenomsystem-transfizierten HEK293-Zellen (3.2.2.2; 3.2.3.2) für nachfolgende RNA-Sequenzierungen (3.2.1.15), qRT-PCR- (3.2.1.3) oder Northern Blot-Analysen (3.2.1.14) wurde mittels des *mirVana™* miRNA Isolation Kits (+ Phenol; TFS) entsprechend der Herstellerempfehlungen durchgeführt. Die RNA-Isolierung beruht auf organischer Extraktion und der Aufreinigung über einen Glasfaserfilter unter Anwendung verschiedener Bindungs- und Waschbedingungen.

Die Aufreinigung zellulärer Gesamt-RNA aus HEK293-Zellen (3.1.6.2) für eine nachfolgende qRT-PCR-Analyse (3.2.1.3) erfolgte mithilfe des RNeasy Mini Kit®s (QIAGEN) entsprechend der empfohlenen Arbeitsanleitung für die RNA-Extraktion aus tierischen Zellen. Die RNA-Isolierung beruht auf der Kombination einer Guanidin-Isocyanat-Zelllyse mit der Aufreinigung über eine Silikamembran, welche mehrere Reinigungs- sowie Trocknungsschritte umfasst. Das Kit eignet sich bevorzugt für die Extraktion von RNAs > 200 Nt (~ 90 % Rückgewinnung/*Recovery*). Dennoch lassen sich ebenfalls gewisse Mengen an RNAs < 200 Nt isolieren (3.2.1.12; Abb. 16). Um mögliche DNA-Kontaminationen zu entfernen, wurde, wie vom Hersteller empfohlen, nach den initialen Nukleinsäureextraktionsschritten ein DNase I-Verdau mit dem RNase-Free DNase Set (QIAGEN), allerdings für 1 h statt 15 min, bei Raumtemperatur auf der Säulenmatrix durchgeführt. Die Elution der RNA erfolgte stets in 42 µL RNase-freiem dH<sub>2</sub>O (Säulentotvolumen ca. 2 µL). Anschließend wurde ein zweiter DNase I-Verdau in der RNA-Lösung durchgeführt.

***DNase I-Verdau in RNA-Lösungen***

Der DNase I-Verdau in Lösung erfolgte unter Verwendung der Ambion™ DNase I (TFS) in einem Gesamtvolumen von 60 µL. Auf Eis wurden die folgenden Komponenten vorsichtig gemischt, herunterzentrifugiert und anschließend bei 37 °C für 1 h inkubiert.

**Tab. 33: Ansatz des DNase I-Verdaus in Lösung**

<b>Komponente</b>	<b>Konzentration</b>
RNA in RNase-freiem dH <sub>2</sub> O (40 µL)	/
10x DNase I Puffer	1x
RiboLock RNase Inhibitor (40 U/ µL)	40 U
Ambion™ DNase I (2 U/µL)	8 U
RNase-freies ddH <sub>2</sub> O	Ad 60 µL

Anschließend wurde die DNase bei 75 °C für 5 min im Heizblock inaktiviert, die Probe mit 40 µL RNase-freiem ddH<sub>2</sub>O auf ein Gesamtvolumen von 100 µL verdünnt und bis zur anschließenden Phenol-Chloroform-Isoamylalkohol-Aufreinigung auf Eis gelagert.

***Phenol-Chloroform-Isoamylalkohol-Aufreinigung und Ethanolfällung***

Nach dem DNase-Verdau wurde die RNA-Lösung 1:1 (d.h. mit 100 µL) Roti®-Phenol/Chloroform/Isoamylalkohol (Roth) versetzt und kräftig für 15 s gevortext. Anschließend erfolgte eine Zentrifugation bei 15 °C und maximaler Drehzahl ( $\geq 13.000$  rpm) für 15 min. Die obere Phase (100 µL) wurde in ein neues Reagiergefäß à 2 mL überführt und mit drei Volumenteilen ( $\approx 300$  µL) eines kalten 30:1 Ethanol (EtOH):Natriumacetat (NaOAc)-Gemisches (3 M NaOAc, pH 5) kräftig gemischt. Nach Zugabe von 1 µL GlycoBlue™ Coprecipitant (TFS) folgte Präzipitation der RNA-Probe bei entweder -80 °C für 15 min oder bei -20 °C für mindestens 1 h (ggf. auch über Nacht). Anschließend wurde die Probe bei 4 °C und maximaler Drehzahl ( $\geq 13.000$  rpm) für mindestens 30 min mittels Zentrifugation pelletiert. Nach Entfernen des Überstandes wurde das RNA-Pellet mit 400 µL 70 % EtOH bei Raumtemperatur für 5 min gewaschen und für 10 min bei 15 °C und maximaler Drehzahl ( $\geq 13.000$  rpm) zentrifugiert. Nach Entfernen des EtOHs trocknete das RNA-Pellet bei 37 °C für 2 – 3 min an der Luft und wurde schließlich in 30 µL RNase-freiem ddH<sub>2</sub>O durch vortexen resuspendiert.

Die Lagerung von RNA erfolgte kurzfristig bei -20 °C und langfristig bei -80 °C. Quantität und Reinheit der RNA wurde am Spektralphotometer Nanodrop ND1000 (Peglab) oder UV5Nano (Mettler Toledo) bestimmt (3.2.1.13). Zusätzlich wurde die Integrität der RNA stichprobenartig per 5 %-PAGE-Analyse überprüft (3.2.1.12).



### **3.2.1.12 Elektrophoretische Auftrennung von Nukleinsäuren**

Die Länge und Reinheit von Nukleinsäuren wurde per Gelelektrophorese überprüft. Es wurden entweder Agarose- oder Polyacrylamid (PAA)-Gele verwendet. Nukleinsäuren sind aufgrund ihres Phosphatrückgrates negativ geladen und wandern daher abhängig von ihrer Größe und Basenzusammensetzung schneller (kleine Fragmente) oder langsamer (große Fragmente) in der Gelmatrix in Richtung Anode. Mithilfe eines Größenstandards (DNA- oder RNA-Marker) kann ihre Größe abgeschätzt werden. Durch Verwendung eines farbigen Nukleinsäure-Interkalators (Ethidiumbromid, bzw. GelRed® Nucleic Acid Gel Stain; Biotium) können die Nukleinsäurebanden unter UV-Licht ( $\lambda = 365 \text{ nm}$ ) sichtbar gemacht werden.

#### **Elektrophoretische Auftrennung von DNA**

Für die elektrophoretische Auftrennung von Plasmid-DNA oder von PCR-Produkten wurden Agarosegele verwendet. Diese setzten sich aus 1x TBE (3.1.4) und 1 – 3 % Agarose GTQ (Roth) zusammen. Zunächst wurde die Agarose in gewünschter Masse eingewogen, anschließend mit 1x TBE versetzt und in einer Mikrowelle geschmolzen. 1x GelRed® (10.000x Stammlösung) wurde hinzugefügt und die Lösung in die entsprechende Gelkammer (mit Taschenkamm) eingefüllt. Die DNA-Proben wurden 1:5 mit 5x DNA-Probenpuffer (3.1.4) gemischt, bevor sie in die Geltaschen pipettiert wurden. Als Laufpuffer wurde 1x TBE verwendet. Die Elektrophorese erfolgte bei 80 – 120 V. Wurde die Agarosegelelektrophorese zu präparativen Zwecken genutzt, so wurden die Banden im Gel nach Detektion unter UV-Licht ( $\lambda = 365 \text{ nm}$ ) mit einem Skalpell herausgeschnitten und die DNA aus dem Gel eluiert (3.2.1.11).

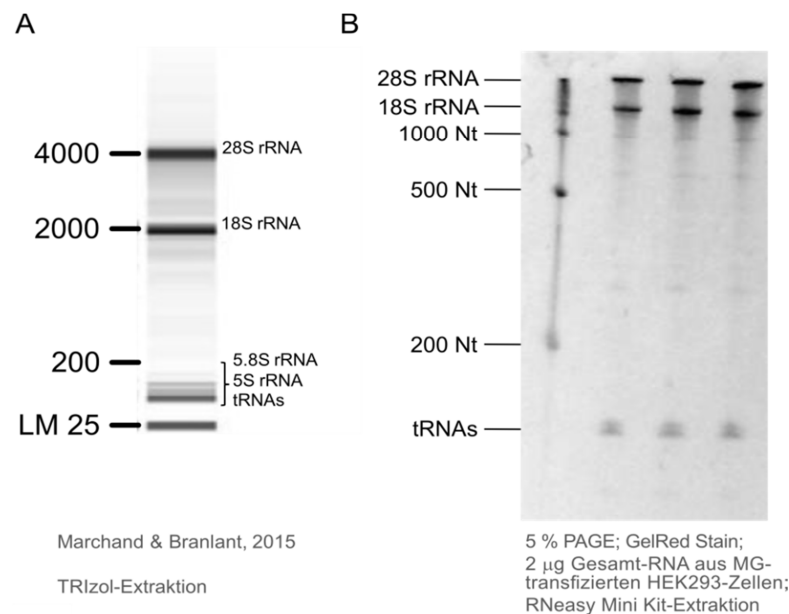
#### **Elektrophoretische Auftrennung von RNA**

Um RNA nach ihrer Extraktion auf ihre Integrität und Reinheit zu überprüfen, wurde eine denaturierende PAA-Gelelektrophorese (PAGE) durchgeführt. Hierfür wurden 2 – 5  $\mu\text{g}$  zellulärer Gesamt-RNA auf ein 5 %-PAA-Gel aufgetragen (Beispiel siehe Abb. 16) und mit GelRed® gefärbt. Distinkte spezifische ribosomale (r)RNA-Banden (v.a. die prominenten 28S und 18S rRNAs bei  $\sim 4000$  und  $\sim 2000 \text{ bp}$ ) sind ein Merkmal für die RNA-Integrität. In denaturierenden PAA-Gelen ist die native RNA-Sekundärstruktur aufgehoben, die RNA liegt einzelsträngig vor. Entsprechend ist das Laufverhalten in der Gelmatrix invers proportional zur Länge der Nukleinsäuresequenz. Eine 50 %-ige PAA-Stammlösung wurde mit einer 8 M Harnstoff-Lösung auf die gewünschte Konzentration verdünnt (3.1.4). Die Initiation der Gelpolymerisation erfolgte mit 0,01 Volumenteilen (VT) des Radikalstarters APS (Ammoniumpersulfat, 10 %) und 0,001 VT des Katalysators TEMED (N,N,N',N'-Tetramethylethylethylendiamin) vor Einfüllen der

Gellösung zwischen zwei Glasplatten. RNA-Proben wurden vor dem Auftragen mit 1x RNA-Probenpuffer (denaturierend) versetzt (3.1.4), für 3 min bei 98 °C erhitzt und sofort auf Eis für 5 min gekühlt. Vor dem Beladen erfolgte eine Spülung der Geltaschen mit 1x TBE (3.1.4), um Harnstoffablagerungen zu entfernen. 1x TBE wurde ebenfalls als Laufpuffer eingesetzt. Die Elektrophorese erfolgte für 5 – 10 %-Gele bei 20 – 35 mA.

### 3.2.1.13 Konzentrationsbestimmung von Nukleinsäuren

Konzentrationen von Nukleinsäurelösungen wurden mittels Spektralphotometrischer Analyse bei einer Welllänge von 260 nm unter Verwendung eines Nanodrop ND1000



**Abb. 16: RNA-Qualitätskontrolle.** Vergleich TRIZOL-Extraktion (A) mit RNeasy Mini Kit-Extraktion (B). **(A)** Analyse der Gesamt-RNA mittels Mikrochip Elektrophoresesystem. LM = Laufpuffermarker. **(B)** Analyse der Gesamt-RNA mittels 5 %-iger denaturierender PAGE und UV-Detektion mit Hilfe des Dark Hood DH-50 Gerix 1000- Geldokumentationssystem. Marker: RiboRuler High Range RNA Ladder (200 - 6.000 bp; Thermo Fisher Scientific).

(Peglab) oder UV5Nano (Mettler Toledo) bestimmt. Die optische Dichte bei 260 nm ( $OD_{260}$ ) ist direkt proportional zum Nukleinsäuregehalt. Für RNA gilt  $A_{260} = 1$  für eine Konzentration von  $\sim 40 \mu\text{g/mL}$ , für doppelstränge DNA gilt  $A_{260} = 1$  für eine Konzentration von  $\sim 50 \mu\text{g/mL}$ . Anhand des Verhältnisses der Absorptionen bei 260 nm und 280 nm kann die Reinheit einer Nukleinsäure-Lösung bestimmt werden. Hier wird unter anderem angenommen, dass Verunreinigungen durch Proteine aufgrund aromatischer Aminosäuren ein Absorptionsmaximum bei 280 nm besitzen. Werte zwischen 1,8 - 2,1 sprechen für reine Nukleinsäurelösungen. Verunreinigungen durch Proteine ergeben einen Quotienten  $< 1,8$ ; Verunreinigungen durch Phenol einen Quotienten  $> 2,0$ .

### 3.2.1.14 Northern Blot-Detektion von RNAs

Die in Manuskript 3 (4.3) dargestellte Northern Blot-Analyse zur Detektion von Leader-Transkripten in Gesamt-RNA aus mit EBOV-infizierten (3.2.3.1) HuH7-Zellen wurde von Dr. Julia Schlereth (Institut für Pharmazeutische Chemie, Marburg) durchgeführt. Die methodische Durchführung ist ihrer Dissertationsschrift vom 14.10.2015 (Schlereth, 2015) sowie dem Methodenteil von Manuskript 3 (4.3) zu entnehmen.

### 3.2.1.15 Nukleinsäuresequenzierung

#### DNA-Sequenzierung

Sequenzierungen von Plasmid-DNA wurden von der Firma SeqLab Microsynth AG (Göttingen) durchgeführt.

Tab. 34: Ansatz der DNA-Sequenzierungen

Bestandteil	Konzentration
Plasmid-DNA (~ 650 ng)	~ 650 ng/15 µL
Sequenzierprimer	2 µM/15 µL
ddH <sub>2</sub> O	Ad 15 µL

Die Auswertung der Sequenzierergergebnisse erfolgte mithilfe der SnapGene-Klonierungsmanager-Software.

#### RNA-Sequenzierung

RNA-Sequenzierungen von EBOV-RNA oder Minigenomsystem-RNA wurden zur Identifikation und Quantifizierung viraler *leader*RNA und mRNA durchgeführt. RNA-Proben < 200 Nt oder Gesamt-RNA wurden von der Firma vertis Biotechnologie AG (Freising) entsprechend analysiert. Vorgehensweisen der cDNA-Bibliothekerstellung und das jeweilige Sequenzierungsprinzip sind dem Material und Methodenteil sowie den Ergänzungen des Manuskripts Nr. 3 (4.3) zu entnehmen. Die bioinformatische Auswertung wurde von Dr. Marcus Lechner (Center for Synthetic Microbiology, Marburg) durchgeführt. Die EBOV-infizierten Zellproben sowie Minigenom-Zellproben für die Sequenzierungen der Jahre 2014 (1439, 1482), 2016 (1647) und 2017 (1757) wurden von Dr. Nadine Biedenkopf (Institut für Virologie, Marburg) generiert.

### 3.2.2 Zellbiologische Methoden

#### 3.2.2.1 Kultivierung von Zellen

##### Kultivierung von Prokaryotischen Zellen auf LB-Agar-Festmedium

Für die Klonierung und Herstellung von Minigenomplasmiden wurden chemisch-kompetente Bakterienzellen der Zelllinie *E. coli* DH5 $\alpha$  (3.1.6.1) verwendet. Die Herstellung des Agar-Festmediums erfolgte wie unter Punkt 3.1.5.1 beschrieben. Bakterienzellen wurden entweder nach Transformation (siehe 3.2.1.10) oder aus Glycerin-Lagerungsstocks (Lagerung bei 80 °C) auf LB-Agarplatten ausplattiert. Ihre Inkubation erfolgte stets über Nacht im Inkubator bei 37 °C. Mit Bakterienkolonien bewachsene Platten wurden im Kühlschrank bei 4 °C bis zur Weiterverwendung gelagert. Für das Inokulieren von Flüssig-LB-Kulturen wurden Einzelkolonien der LB-Agar-Platten verwendet.

##### Kultivierung von Prokaryotischen Zellen in LB-Flüssigmedium

Die Kultivierung von Bakterienzellen in Flüssigmedium erfolgte unter mikrobiellen Standardbedingungen bei 37 °C unter permanenter Sauerstoffdurchmischung im Schüttelinkubator bei 180 – 200 rpm.

##### Kultivierung von HEK293-Zellen

Im Rahmen der Minigenomassays wurde die Zelllinie HEK293 verwendet, bei der es sich um menschliche embryonale Nierenzellen handelt. Ihre Kultivierung erfolgte in 75 cm<sup>2</sup> Zellkulturflaschen im Brutschrank bei 37 °C, 5 % CO<sub>2</sub> und in feuchter Atmosphäre. Als Zellkulturmedium wurde DMEM (3.1.5.2; Thermo Fisher Scientific) verwendet, welches mit 50 U/mL Penicillin, 50 µg/mL Streptomycin, 2 mM L-Glutamin und 10 % fetalem Kälberserum (FCS) supplementiert wurde (DMEM<sub>10%FCS+Q+P/S</sub>). Eine Passagierung der Zelllinie erfolgte alle drei bis vier Tage. Der konfluente Zellrasen wurde nach Mediumabnahme zunächst zweimal mit circa 10 mL des zuvor im Wasserbad auf 37 °C erwärmten PBS<sub>def.</sub>-Puffers (3.1.4) gewaschen. Durch Zugabe von 2 mL eines Trypsin/EDTA (0,5 %)-Gemisches wurden die Zellen anschließend vom Flaschenboden abgelöst. Nach mikroskopischer Überprüfung der vollständigen Ablösung wurde die Reaktion mittels Zugabe von 8 mL des Vollmediums (DMEM<sub>10%FCS+Q+P/S</sub>) gestoppt. Die Zellen wurden in der gewünschten Verdünnung im Medium resuspendiert und anschließend in eine neue Zellkulturflasche oder für anschließende Minigenomassays (3.2.3.2) in 6-Well-Zellkulturplatten ausgesät.

### 3.2.2.2 *Transiente DNA-Transfektion mit TransIT®-LT1*

In der vorliegenden Arbeit wurde Plasmid-DNA im Rahmen des EBOV-spezifischen Minigenomsystems (3.2.3.2) über transiente Transfektion mit TransIT®-LT1 (mirus) in eukaryotische Zellen eingebracht. Das TransIT®-LT1 besteht aus Lipidpolymeren, welche mit der Plasmid-DNA über die Anlagerung an das negativ geladene DNA-Phosphatrückgrat Lipid-DNA-Komplexe bilden. Diese Lipid-DNA-Komplexe können sich effizient an Zellmembranen anlagern und diese überwinden. Die Transfektion von Plasmid-DNA in HEK293-Zellen mit TransIT®-LT1 erfolgte stets 18 – 24 h nach ihrer Aussaat bei einer Zellkonfluenz von 60 - 80% in Wachstums-Vollmedium (DMEM+10%FCS+P/S+Q; 3.1.5.2), sofern nicht anders beschrieben. Es wurde, gemäß der Herstellerempfehlung, kein Medienwechsel im Anschluss an die Transfektion durchgeführt. Zunächst wurden 100 µL Opti-MEM® I Reduced-Serum Medium (TFS) mit 6 µL TransIT®-LT1-Reagenz (3 µL TransIT® pro 1 µg Plasmid-DNA) versetzt und für 5 min inkubiert. Während der Inkubationszeit wurden zudem in einem separaten Mikroschraubröhrchen 100 µL Opti-MEM® mit der Plasmid-DNA (insgesamt 1950 ng) gemischt. Nach 5-minütiger Inkubationszeit wurde der TransIT-LT1:Opti-MEM-Reaktionsansatz mit dem Plasmid-DNA:Opti-MEM-Reaktionsansatz vereinigt, vorsichtig gevortext und anschließend bei Raumtemperatur für 15 – 30 min inkubiert. Abschließend wurden die gebildeten TransIT-LT1:Plasmid-DNA-Komplexe auf die HEK293-Zellen geträufelt und diese bei 37 °C im Brutschrank bei 5 % CO<sub>2</sub>-Begasung für 48 h kultiviert (3.2.2.1).

### 3.2.2.3 *Ernte von HEK293 Zellen*

Die Ernte von HEK293-Zellen nach Minigenomsystem-Assay (3.2.3.2) erfolgte 48 h nach Transfektion (3.2.2.2). Hierbei wurden die Zellen zunächst einmal mit 1 mL PBS<sub>def.</sub>/well gewaschen, anschließend mit dem Zellschaber abgekratzt und jeweils 500 µL der PBS-Zellsuspension in ein Reagiergefäß überführt. Die Auftrennung auf zwei Reagiergefäße ermöglichte die separate Zelllyse einerseits für nachfolgende Luziferase-Assays (3.2.2.4) und andererseits für nachfolgende RNA-Extraktionen zwecks qRT-PCR-Analyse (3.2.1.11; 3.2.1.3). Die Zellen aus beiden Reagiergefäßen wurden 2 min bei 3000 rpm pelletiert und der Überstand verworfen. Im Anschluss wurde das Pellet des einen Reagiergefäßes in 200 µL 1x Lysis-Juice (PJK) und das Pellet des zweiten Gefäßes in 350 µL Lysepuffer des QIAGEN RNeasy Mini Kit®s resuspendiert. Letztgenannte Lysate wurden bis zur RNA-Extraktion bei -80 °C gelagert. Erstgenannte Lysate wurden entweder für 1 h inkubiert und zwischendurch gevortext oder über Nacht bei -20 °C inkubiert und anschließend gevortext. Es folgte eine Zentrifugation der Lysate bei 4 °C

und 14.000 rpm für 5 min. Die Proben wurden entweder direkt im Luziferase-Assay vermessen oder bei -20 °C bis zur späteren Messung gelagert.

#### **3.2.2.4 Luziferase-Assay**

Für die Messung der Renilla (Rluc)-Luziferaseaktivität aus Replikations-kompetenten (RC) monocistronischen Minigenomassays (MG-Assays) wurden die Lysate des Zentrifugationsüberstandes aus 3.2.2.3 zunächst 1:50 mit dH<sub>2</sub>O verdünnt. Firefly (Fluc)-Luziferaseaktivitäten wurden stets aus unverdünnten Lysaten bestimmt. Selbiges gilt sowohl für Rluc- als auch Fluc-Messungen von Replikations-defizienten (RD) oder bicistronischen MG-Assays. 10 µL der verdünnten oder unverdünnten Lysate wurden in ein Well einer 96-well-Microtiterplatte überführt. Die Messung der Luziferaseaktivität erfolgte unter Zugabe von 50 µL des jeweiligen Substratpuffers Firefly-Luziferase-Reagenz (Beetle-Juice; PJK) oder Renilla-Luziferase-Reagenz (Renilla-Juice Fluid; PJK) 0,5 s (Renilla) bzw. 5 s (Firefly) nach Substratpufferzugabe in einem Centro LB 960 Luminometer (Berthold). Im Falle monocistronischer MG-Assays (RC oder RD) wurden die Rluc-Aktivitäten jeweils durch Division mit den korrespondierenden Fluc-Aktivitäten normalisiert, um Transfektionseffizienzunterschiede zu berücksichtigen. Alle auf diese Weise normalisierten Aktivitätswerte wurden anschließend auf das Wildtyp (Wt)-MG normalisiert (Wt-Aktivität wurde auf 100 % gesetzt). Relative Rluc- oder Fluc-Aktivitätswerte aus bicistronischen MG-Assay-Proben wurden durch jeweilige Normalisierung auf den Rluc- oder Fluc-Aktivitätsabsolutwert des bicistronischen Referenz-MGs (auf 100 % gesetzt) generiert.

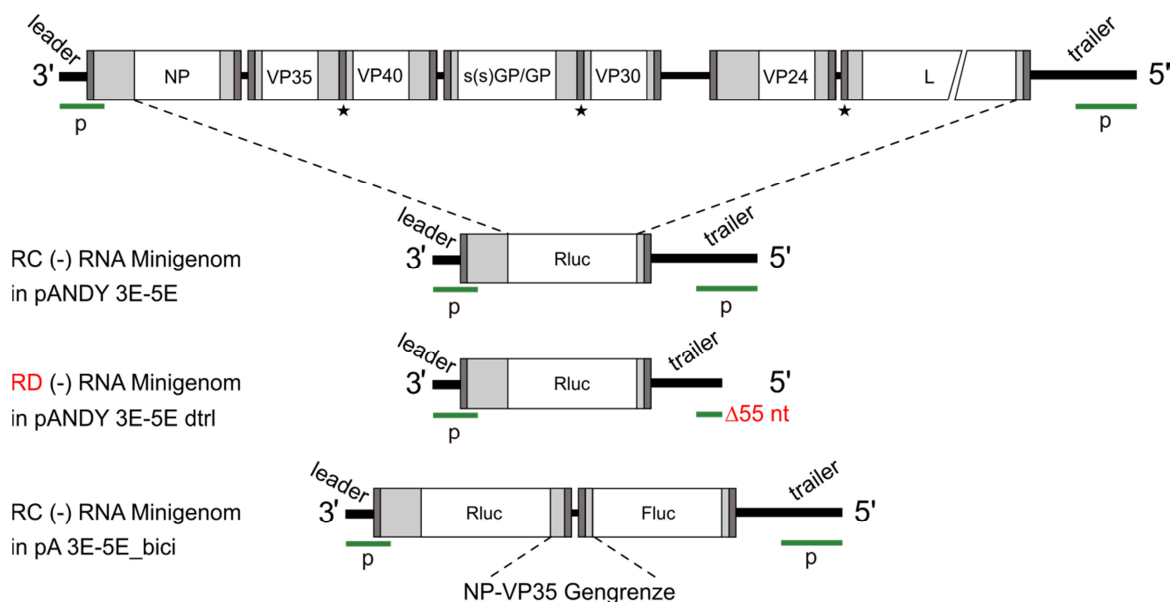
### 3.2.3 Virologische Methoden

#### 3.2.3.1 Ebola-Virus-Infektion von HuH7-Zellen

Die Infektion von HuH7-Zellen mit EBOV (3.1.7) erfolgte im BSL-4-Labor des Instituts für Virologie in Marburg durch Dr. Nadine Biedenkopf und ist dem Methodenteil von Manuskript 3 zu entnehmen (4.3). Anschließende Gesamt-RNA-Extraktionen der BSL-4-Proben (3.2.1.11) wurden ebenfalls von Dr. Nadine Biedenkopf durchgeführt.

#### 3.2.3.2 Ebola-Virus-spezifisches Minigenomsystem

Das Ebola-Virus (EBOV)-spezifische Minigenomsystem wurde analog zu 1.9, Abb. 13a) verwendet. Alle eingesetzten Minigenome (MG) basieren auf einem der folgenden drei Minigenomvarianten:



**Abb. 17: EBOV-spezifische Minigenomsysteme.** Das Farbschema ist grundsätzlich wie in Abb. 8. Im Minigenom (MG)-System sind die offenen Leserahmen (ORFs) der EBOV Proteine durch ein oder mehrere Reporterproteingen(e) [Renilla (Rluc)- und/oder Firefly (Fluc)-Luziferase], ersetzt. Der regulatorische 3'-Leader und der 5'-Trailer bleiben enthalten. Im Falle der monocistronischen (1 ORF) MG (Replikationskompetent; RC und Replikations-defizient; RD) bleiben zudem die NP-5'-UTR mit dem ersten NP-Genstart (GS)-Signal sowie die L-3'-UTR mit dem L-Genend (GE)-Signal intakt. Im bicistronischen MG (2 ORFs; pA 3E5E\_bici/pA 3E5E\_Rluc\_Fluc) ist darüber hinaus die authentische NP-VP35-Gengrenze enthalten. Im RD-Minigenom sind die 5'-terminalen 55 Nt deletiert. Dies inaktiviert den Trailer-Replikationspromoter. Somit kann keine vRNA-Synthese vom cRNA-Templat stattfinden (single-round Replikation/Transkription).

Das EBOV-spezifische Minigenom kodiert für eine Renilla-Luziferase (Rluc) oder, im Falle des bicistronischen MGs, für eine Rluc und eine Firefly-Luziferase (Fluc). Es steht unter der Kontrolle eines T7-Promotors und wird in der transfizierten Zelle (3.2.2.2) von der auf dem pCAGGS-T7-Plasmid koexprimierten T7-Polymerase in eine Negativstrang-(Mini-)Genom-RNA transkribiert. Diese ist das erste Templat für die virale Pirmärtranskription in der transfizierten Zelle. Alle pCAGGS-Plasmide sind unter der

Kontrolle eines CAG-Promotors, sie werden daher von Wirtszellpolymerasen überexprimiert. Sie liefern neben der T7-Polymerase die Transkriptions- und Replikations-relevanten Nukleokapsidkomponenten NP, VP35, L und VP30. Falls nicht anders beschrieben, wurden die nachfolgenden tabellarisch gelisteten Plasmidkonzentrationen mittels mit TransIT®-LT1 in  $8 \times 10^5$  HEK293-Zellen pro Well einer 6-Well-Zellkulturplatte transfiziert (3.2.2.2):

Tab. 35: Pipettierschema Minigenom-Assay

Plasmid-DNA	Konzentration [ng]
pCAGGS-NP	125
pCAGGS-VP35	125
pCAGGS-VP30	100
pCAGGS-L	1000
pCAGGS-T7	250
Minigenom (Abb. 15):	
- pANDY 3E-5E	250
- pANDY 3E-5E dtrl	
- pANDY 3E-5E_bici	
pGL4.13 [luc2/SV40]	100

Die auf dem pGL4.13-Plasmid kodierte Firefly-Luziferase dient der späteren Normalisierung der Rluc-Biolumineszenzwerte im Luziferase-Assay (3.2.2.4) des jeweiligen monocistronischen Minigenoms (pANDY 3E-5E, pANDY 3E-5E dtrl oder Derivaten davon). Im Falle der bicistronischen Minigenome [pANDY 3E-5E\_bici und Derivate (siehe Kapitel 4.3)] ist bereits eine Firefly-Luziferase im zweiten Cistron des Genoms kodiert, weshalb die Normalisierung nicht über eine kotransfizierte Firefly-Luziferase erfolgen kann. Anstelle des pGL4.13-Plasmids wurde in diesem Fall der pCAGGS-Leervektor in entsprechender Menge transfiziert, um die Gesamtmenge der Plasmid-DNA (1950 ng) beizubehalten. Gleiches gilt auch für Minigenomsysteme unter Ausschluss von VP30 oder L. Hier wurde das Fehlen des pCAGGS-VP30- oder pCAGGS-L-Plasmids während der Transfektion ebenfalls durch den pCAGGS-Leervektor ausgeglichen.



## 4 ERGEBNISSE UND DISKUSSION

### 4.1 Die Rolle der „Rule of 6“ für die EBOV-Transkription

Der genomische Ebola-Virus (EBOV)-Replikationspromotor ist in die zwei Promotorelemente PE1 und PE2 unterteilt. PE1 umfasst die 55 3'-terminalen Nukleotide des Genoms und wird durch die Transkriptionsstartsequenz (TSS, 12 Nt) des ersten Nukleoprotein (NP)-Gens sowie eine Spacer-Sequenz (13 Nt) von PE2 getrennt. In einer früheren Studie von Weik *et al.* (2005) konnte gezeigt werden, dass Insertionen oder Deletionen innerhalb des Spacers nur dann eine produktive virale Replikation ermöglichen, wenn sie 6 oder 12, nicht jedoch 1/2/3/5 oder 9 Nukleotide umfassen. Dies veranlasste Weik *et al.* zur Formulierung der „Rule of 6“ für den EBOV-Replikationspromotor. In der vorliegenden Studie haben wir untersucht, ob das Hexamerraster ebenfalls die virale Transkription beeinflusst. Hierbei kamen eine Reihe Replikations-kompetenter und -defizienter monocistronischer Minigenome zum Einsatz, in denen der Spacer des NP-Gens entweder mutiert oder durch Genstart-Bereiche interner EBOV-Gene sowie mutierter Varianten dieser ersetzt wurde. Neben Reporter-gen-Assays verwendeten wir qRT-PCRs zur Quantifizierung viraler mRNA, genomischer vRNA und antigenomischer cRNA. Wir konnten so zeigen, dass das Hexamerraster sowohl für die Replikation als auch die virale Transkription essenziell ist. Darüber hinaus legen unsere Ergebnisse den Schluss nah, dass sich die 3'-UN<sub>5</sub>-Hexamerperiodizität von PE2 über den TSS-Spacer-Bereich bis in den 5'-terminalen Bereich von PE1 fortsetzt. Die Spacer-Region kann ohne Transkriptionsaktivitätsverluste um 48 Nt verlängert werden. Ein kontinuierliches UN<sub>5</sub>-Hexamerraster zwischen PE1 und PE2 steigert sowohl die Transkription als auch die Replikation. Wir haben ein Modell zur Diskussion gestellt, nach dem das UN<sub>5</sub>-Hexamerraster daran beteiligt sein könnte, NP für die Initiation der viralen RNA-Synthese in einem definierten Raster auf dem RNA-Templat zu assemblieren, um so der RNA-Polymerase eine produktive Templat-Bindung zu ermöglichen. Ebenfalls denkbar wäre eine Rolle in der koordinierten Dissoziation und Reassoziierung von NP-Proteinen an der Templat-RNA, wenn diese während der Elongationsphase der Replikation oder Transkription durch das aktive Zentrum des Polymerasekomplexes gefädelt wird. Unsere Daten weisen zudem darauf hin, dass PE1 den Transkriptionspromotor enthält, da eine Deletion der 5'-terminalen PE1-Nukleotide -54 und -55 die virale Transkription vollständig zum Erliegen bringt.

## **Hexamer phasing governs transcription initiation in the 3'-leader of Ebola virus**

Simone Bach<sup>1</sup>, Nadine Biedenkopf<sup>2</sup>, Arnold Grünweller<sup>1</sup>, Stephan Becker<sup>2</sup>, Roland K. Hartmann<sup>1, #</sup>

<sup>1</sup>Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35037 Marburg, Germany

<sup>2</sup>Institut für Virologie, Philipps-Universität Marburg, Hans-Meerwein-Str. 2, 35043 Marburg

#corresponding author:

E-mail:       roland.hartmann@staff.uni-marburg.de

Running head: hexamer phasing in EBOV transcription

Keywords: viral transcription, transcription promoter, minigenome systems, 3'-leader mutants, 3'-UNNNNN hexamer phasing

---

**ABSTRACT**

The genomic, bipartite replication promoter of Ebola virus (EBOV) consists of elements 1 (PE1) and 2 (PE2). PE1 (55 nt at the 3'-terminus) is separated from PE2 (harboring eight 3'-UN<sub>5</sub> hexamers) by the transcription start sequence (TSS) of the first nucleoprotein (NP) gene plus a spacer sequence. Insertions or deletions in the spacer were reported to support genome replication if comprising 6 or 12, but not 1/2/3/5/9 nucleotides. This gave rise to formulation of the "rule of 6" for the EBOV replication promoter. Here, we studied the impact of such hexamer phasing on viral transcription using a series of replication-competent and -deficient monocistronic minigenomes, in which the spacer of the NP gene was mutated or replaced with that of internal EBOV genes and mutated variants thereof. Beyond reporter gene assays, we conducted qRT-PCR to determine the levels of mRNA, genomic and antigenomic RNA. We demonstrate that hexamer phasing is also essential for viral transcription, that UN<sub>5</sub> hexamer periodicity extends into PE1 and that the spacer region can be expanded by 48 nt without losses of transcriptional activity. Making the UN<sub>5</sub> hexamer phasing continuous between PE1 and PE2 enhanced the efficiency of transcription and replication. We show that the two nucleotides preceding the TSS are essential for transcription. We further propose a role for UN<sub>5</sub> hexamer phasing in positioning NP during initiation of RNA synthesis, or in dissociation/reassociation of NP from the template RNA strand while threading the RNA through the active site of the elongating polymerase during replication and transcription.

## INTRODUCTION

The family *Filoviridae* in the order *Mononegavirales* (Bukreyev et al., 2014) includes five genera (Cuevavirus, Ebolavirus, Marburgvirus, Striavirus, and Thamnovirus) and a total of nine species. Members of this family form variously shaped, often filamentous, enveloped virions containing linear non-segmented, negative-sense (NNS) RNA genomes of 15 – 19 kb (Kuhn et al., 2019b).

Filoviruses are causative agents of severe hemorrhagic fever in humans and non-human primates with reported case fatality rates between 34 and 81 % in humans (Burk et al., 2016). The 2014/15 outbreak in West Africa and the current EBOV epidemic in the Democratic Republic of Congo suggest that filoviruses have become a constant threat of public health.

The negative-sense RNA genome of EBOV has a length of 19 kb with regulatory regions at the 5'- and 3'-end (termed 5'-trailer and 3'-leader) that harbor replication promoters and, specific for the 3'-leader, the first transcription start sequence (TSS) for the synthesis of the nucleoprotein (NP) mRNA (Mühlberger, 2007; Sanchez et al., 1993). The helical viral nucleocapsid, consisting of the viral genome in complex with the viral proteins NP, VP35 and VP24, serves as the template for replication and transcription. Recent cryo-EM structures of NP:RNA complexes verified previous predictions that 6 RNA residues are bound per NP monomer (Sugita et al., 2018; Wan et al., 2017). For viral transcription, the polymerase L and its cofactor VP35 require association with the EBOV transcription factor VP30 (Mühlberger et al., 1999). Evidence obtained for the Mumps virus phosphoprotein (P) (Cox et al., 2014), the putative ortholog of VP35, suggests that VP35 may target L to the nucleocapsid and unwind the latter to allow the polymerase to gain access to the genomic RNA. Additionally, a novel NTP-ase and helicase function was revealed for VP35, with a preference for unwinding of dsRNA helices with 5' overhangs (Shu et al., 2019). Inhibition of helicase activity resulted in a decrease of viral transcription and replication in cells, suggesting that RNA secondary structures need to be resolved during viral transcription/replication. Transcription is positively influenced by interaction of VP35 and VP30 (Biedenkopf et al., 2013) which was shown to be mediated by RNA and to depend on the dsRNA-binding capacity of VP35 (Biedenkopf et al., 2016b).

The phosphoprotein VP30 preferentially binds single-stranded RNA (ssRNA) of mixed base composition in conjunction with stem-loop structures (Schlereth et al., 2016). VP30 supports transcription in its non-phosphorylated state, whereas VP30 phosphorylation particularly at the key serine 29 inhibits viral transcription (Biedenkopf et al., 2016a;

Kruse et al., 2018; Modrof et al., 2002). Phosphorylation of VP30 strengthens its interaction with NP and weakens the interaction with the polymerase cofactor VP35, which is thought to result in release of VP30 from the viral transcription complex (Biedenkopf et al., 2013). There is evidence that multiple phosphorylation/dephosphorylation events mediate the protein's dynamic association and dissociation with the nucleocapsid and transcription complexes throughout the viral life cycle (Biedenkopf et al., 2016a).

The EBOV replication promoters could be confined to the 3'-terminal 156 nt of the genomic leader and the 3'-terminal 177 nt of the antigenomic RNA (trailer promoter) (Calain et al., 1999). In this context, it is worth mentioning that cRNA synthesis is initiated opposite to the penultimate C residue of the genome's 3'-end in single-round replicating minigenomes (Deflubé et al., 2019).

Mühlberger and coworkers (Weik et al., 2005) performed a mutational investigation of the genomic 3'-leader promoter in a replication-competent minigenome system. They concluded that the EBOV replication promoter is bipartite (Fig. 1B): promoter element 1 (PE1) comprises the 3'-terminal 55 nucleotides of the genomic leader up to the TSS directing transcription of the first NP mRNA; promoter element 2 (PE2) comprises eight consecutive 3'-UN<sub>5</sub> hexamers in the region of nt -81 to -128. Mutational analysis of UN<sub>5</sub> hexamers suggested that the five 5'-proximal UN<sub>5</sub> hexamers in PE2 are sufficient for normal replication activity and the 5'-proximal three for residual replication activity (Weik et al., 2005). Finally, deletions and insertions of up to 12 nt in the spacer region between TSS and PE2 (Fig. 1B) revealed that only insertions or deletions of 6 or 12 nt, but not 1, 2, 3, 5 or 9 nt, are compatible with replication (Weik et al., 2005). This motivated Weik et al. (2005) to propose the "rule of 6" for the EBOV replication promoter, although the EBOV genome is not divisible by 6. The term "rule of 6" was originally coined in the paramyxovirus field to indicate the dependency of viral polymerase activity on hexamer phasing of the entire genome (Calain and Roux, 1993; Kolakofsky et al., 1998). To avoid terminological confusion, we have preferentially used the term "hexamer phasing" instead of "rule of 6" for the EBOV system.

However, the reference point for hexamer phasing in the promoter region remained unclear in the study by Weik et al. (2005), and the results left open whether hexamer phasing also pertains to viral transcription. In the present study, we sought to define such a reference point for hexamer phasing, how this pattern may functionally link PE1 and PE2 and if there are additional length, sequence and structural constraints for the region connecting PE1 and PE2. We addressed these questions by investigating a series of

replication-competent and -deficient monocistronic minigenomes (MGs), in which the terminal part of the 5'-UTR coding region of the NP gene was mutated or replaced with that of internal EBOV genes and mutated variants thereof. In addition to reporter gene assays as a measure of mRNA synthesis by the EBOV polymerase complex, we conducted qRT-PCR to determine the levels of mRNA, genomic RNA (abbreviated as vRNA for viral RNA) and antigenomic RNA (abbreviated as cRNA for complementary or copy RNA) to assess effects on transcription and replication.

## RESULTS

The EBOV genome and the potential RNA secondary structures harboring the seven EBOV transcription start sequences (TSS; highlighted in cyan) are depicted in Fig. 1A. The current state of knowledge regarding PE1 and PE2 of the 3'-leader, including the eight 3'-UN<sub>5</sub> hexamers in PE2, the TSS, the spacer sequence (in orange) and experimentally verified secondary structures (Schlereth et al., 2016; Weik et al., 2005) of the (naked) genomic RNA and its antigenomic copy are summarized in Fig. 1B.

### Replacing the NP hairpin structure with hairpin structures of internal EBOV genes

As a first approach to study the sequence, structure and spacing constraints for replication and transcription in the region between PE1 and PE2, we replaced the hairpin (HP) structure harboring the TSS of the first NP gene with the corresponding HP structures of internal EBOV genes in the replication-competent (RC) monocistronic MG (Fig. 2A). Remarkably, only the MG with the VP40 HP substitution gave rise to wild type-like reporter activity, whereas constructs with HP structures derived from the VP35, GP, VP30, VP24 and L genes failed to do so (Fig. 2B). When reinspecting the sequence of the 3'-leader, we noticed that the pattern of 3'-UN<sub>5</sub> hexamers (termed UN<sub>5</sub> hexamers in the following) could be extended from PE2 to position -51 in PE1 with a single interruption at G<sub>-75</sub> (Fig. 2C). This perspective suggests that hexamer periodicity may already begin in PE1 and thus may functionally link PE1 and PE2. Indeed, the distance of U<sub>-51</sub> to position -80 has a length of 30 nt in case of the wild type (wt) NP hairpin (Fig. 2C). In the construct with the VP40 hairpin, this distance is a multiple of 6, as well (66 nt), but in all other aforementioned constructs the distance deviates from this hexamer phase (Fig. 2C). This finding made it plausible to define U<sub>-51</sub> as a reference point for UN<sub>5</sub> hexamer phasing in the 3'-leader.

## Engineering of hairpins to comply with hexamer phasing

We next examined if the inactive constructs could be reactivated by adjusting the HP structures to a length that conforms to hexamer phasing (mutant constructs are illustrated in Fig. 3). Inserting 1 nt in the stem or loop of the VP35 HP to elongate the distance between U<sub>-51</sub> and the first UN<sub>5</sub> hexamer in PE2 from 29 to 30 nt indeed fully restored reporter gene activity (Fig. 4A). Basically the same pertains to the engineered VP30 HP (insertion of 4 nt in the stem), the L HP (1 nt insertion in the stem or loop) and the GP HP (deletion of 2 nt in the loop; restoration of reporter activity to ~ 50 % of wt). In contrast, deletion of nt -54 and -55 preceding the GP stem (Fig. 3) failed to restore activity (Fig. 4A, construct GP-2 ( $\Delta$ 2 nt PE1)), indicating inactivation of the promoter by deletion of nucleotides in PE1. This is in line with the finding that mutation of nucleotides -53 to -55 abolished the production of replicated RNA (Weik et al., 2005). For the VP24 HP construct, 2 nt insertions in the stem or loop to adjust hexamer phasing only weakly restored reporter activity. As the engineered VP24 HP expands the spacer region by 54 nt relative to the wt NP HP, this result may be attributable to length limitations for the distance between PE1 and PE2. The results shown in Fig. 4A support the concept of using position -51 in PE1 as reference point for hexamer phasing in the 3'-leader promoter (Fig. 2C), which functionally links PE1 to PE2. In addition, the results demonstrate considerable tolerance for variation in the spacer sequence between PE1 and PE2 in terms of sequence (except for the conserved TSS), length and structure as long as hexamer phasing is maintained.

We further pursued a reverse engineering approach, where we inserted or deleted a single nucleotide in the native NP HP (Fig. 4B). This was done to exclude the influence of sequence and length variation that might have contributed to the functional readout of the constructs in which the wt NP HP was replaced with hairpins of internal EBOV genes. This 1-nt insertion or deletion, either in the stem or loop of the NP HP, abolished reporter activity, verifying that conformance with hexamer phasing is a basic prerequisite for reporter activity.

## Hexamer phasing in the context of replication-deficient minigenomes

In reverse-genetics systems using replication-competent MGs, the amplitude of reporter gene expression is not only a measure of viral transcription, but also includes contributions from viral replication that leads to the synthesis of new vRNA nucleocapsids that are efficient templates for transcription (Hoenen et al., 2010). To provide further evidence that hexamer phasing also applies to transcription, we tested the constructs described above in the context of the replication-deficient (RD) MGs

(Hoenen et al., 2010). In the RD MG system, where only transcription takes place, viral mRNA levels are much lower than in the RC MG system, most likely due to low-efficiency assembly of functional nucleocapsids from minigenomic vRNAs produced by the T7 RNA polymerases and coexpressed NP protein in the absence of viral replication (Blakqori et al., 2003). Only replicative synthesis of new vRNA by the viral polymerase in the RC MG system will direct the assembly of functionally more competent nucleocapsids that are much more efficient templates for viral transcription. We tried to quantify the viral mRNA levels for the RD NP (wt) MG construct by qRT-PCR, but could not determine viral mRNA-specific  $C_T$  values sufficiently above background (-L control): this might be attributable to T7 RNA polymerase also synthesizing some antiminigenomic RNAs that are recognized by the RT primers specific for the viral RNAs (cRNA, mRNA). Comparable experiences were made in the previous study (Hoenen et al., 2010). Nevertheless, viral mRNA levels in the RD MG system were sufficient for reporter enzyme activity measurements reliably above background. As a result, we observed the same picture in the RD versus RC MG setup: the VP40 construct gave rise to NP (wt)-like reporter activity (and, by inference, to viral mRNA transcription), whereas other replacement constructs, such as those harboring the VP35, GP or VP30 HP, failed to do so (Fig. 5A). Likewise, engineered HP constructs designed to restore hexamer periodicity also restored reporter activity to similar relative extents as in the replication-competent system, as exemplified for the VP35, GP and VP30 hairpins (cf. Fig. 4A and 5B). Conversely, 1-nt deletions or insertions in the wt NP HP abolished reporter activity (Fig. 5C). As reporter activities in the RD MG system depend on viral transcription in the absence of replicative vRNA synthesis, and assuming that translation efficiencies of the tested mRNA variants as well as their half-lives were comparable, our findings indicate, for the first time, that hexamer phasing is mandatory for EBOV transcription.

### **Residual polymerase activity upon disruption of hexamer phasing**

Using qRT-PCR, we quantified the levels of mRNA, cRNA and vRNA in the replication-competent MG system for the wt, VP35, VP35+1 (loop), NP-1 (stem) and the GP-2 ( $\Delta 2$  nt PE1) constructs relative to the L omission control. The reporter activities measured for these variants in the RC and RD MG systems, already shown in Fig. 4 and 5, were included in Fig. 6A, B for immediate comparison with the qRT-PCR data. Individual constructs showed very similar relative differences in reporter activity in the RD and RC MG setups (Fig. 6A, B). However, the absolute amounts of viral mRNA were 2-3 orders of magnitude higher (this study; Hoenen et al., 2010) in the RC versus RD MG system as discussed above, indicating that viral replication yields new copies of genomic vRNA, each of which is efficiently transcribed multiple times. Thus, viral



replication boosts transcription, although the majority of RNA synthesis products are viral mRNAs as illustrated by comparison of  $2^{-\Delta CT}$  values for mRNA vs. cRNA and vRNA in the RC MG system (Fig. 6C). The relative levels of mRNA, which could only be reliably quantified in the RC system (Fig. 6D), matched the pattern seen for reporter activity: mRNA levels in cells transfected with the VP35, NP-1 (stem) or GP-2 ( $\Delta 2$  nt PE1) construct were as low [GP-2 ( $\Delta 2$  nt PE1)] or almost as low as in the background control (L omission), whereas the VP35+1 (loop) construct, reengineered to conform to hexamer phasing, restored wt-like mRNA levels (Fig. 6D). The reduction of mRNA levels essentially to background levels in the case of the GP-2 ( $\Delta 2$  nt PE1) construct reveals that deletion of nt -54 and -55 in PE1 essentially prevents any viral transcription activity in the 3'-leader. In turn, mRNA levels above background for the other two constructs with disrupted hexamer phasing [VP35, NP-1 (stem)] suggests residual activity of the viral RNA polymerase despite non-compliance with hexamer phasing. However, the extent of this hexamer phasing-independent transcription seems to be context-dependent; this is suggested by higher mRNA levels for the VP35 versus NP-1 (stem) variant (Fig. 6D).

Among the two variants with disrupted hexamer phasing, cRNA and vRNA levels, after subtraction of background (-L control), were higher for the VP35 than for the NP-1 (stem) variant (Fig. 6E, F), in line with the corresponding hierarchy of mRNA levels (Fig. 6D). Surprisingly, particularly the vRNA levels for the GP-2 ( $\Delta 2$  nt PE1) variant were substantial (Fig. 6F), in contrast to the essentially insignificant mRNA levels (Fig. 6D). This may suggest that the deletion of nt -54/55 impairs transcription more severely than replication.

In summary, the results presented in Fig. 6C illustrate that the majority of RNA products generated by the viral RNA polymerase in the RC MG system are mRNAs and viral replication boosts transcription. The data in Fig. 6C suggest that each vRNA nucleocapsid newly synthesized by the viral polymerase serves as template for  $\sim 13$  rounds of mRNA synthesis. Polymerization activity is highly sensitive to violation of hexamer phasing, but residual and context-dependent polymerase activity still occurs upon deviation from the hexamer phase; a dinucleotide deletion in PE1, immediately preceding the TSS [construct GP-2 ( $\Delta 2$  nt PE1)], essentially abolished transcription.

### **Extension of the UN<sub>5</sub> hexamer pattern**

We next addressed the question whether extension of the UN<sub>5</sub> hexamer pattern between PE1 and PE2 may influence reporter activity. UN<sub>5</sub> hexamer phasing is continuous from nt -51 in PE1 to PE2 with a single interruption at G<sub>-75</sub> (Fig. 2C). We thus mutated

G<sub>-75</sub> to U, resulting in variant NP U<sub>-75</sub>. As RNAfold and mfold predicted almost complete disruption of secondary structure for this variant, we constructed variant NP U<sub>-75</sub>/G<sub>-72</sub>, in which a second mutation (G<sub>-72</sub>) introduced a G:C bp into the upper stem region to compensate for the destabilizing effect of the U<sub>-75</sub> mutation (Fig. 7A). The hairpin stability of variant NP U<sub>-75</sub>/G<sub>-72</sub> is predicted to be similar to that of the wt NP hairpin. We further constructed variant NP G<sub>-72</sub> carrying only the G<sub>-72</sub> mutation (Fig. 7A) to control for stability effects. Variants NP U<sub>-75</sub> and NP U<sub>-75</sub>/G<sub>-72</sub> with continuous UN<sub>5</sub> hexamer phasing gave rise to increased reporter activity in RC and RD MG systems relative to the NP and NP G<sub>-72</sub> controls (Fig. 7B). This correlated with increased mRNA and cRNA levels in the RC system; corresponding increases in the vRNA level were less pronounced (Fig. 7C). The latter result might be attributable to the narrower dynamic range of variations in the amount of vRNAs that are pre-synthesized by the T7 RNA polymerase in the MG system. Relative increases in reporter activity for the U<sub>-75</sub> variants were particularly pronounced in the RD MG system, again consistent with the notion that continuous UN<sub>5</sub> hexamer phasing strongly impacts viral transcription.

We conclude that increasing the number of UN<sub>5</sub> hexamers between nt -51 and -80 can enhance transcriptional but also replicative activity of the viral polymerase.

## DISCUSSION

The 3'-leader regions of NNS RNA viruses harbor essential information for viral RNA synthesis: promoters that allow either transcription of viral mRNAs or enable replication into full-length genomes (vRNA) via antigenomes (cRNA) as intermediates. Genomic promoter architectures of NNS viruses were divided into two groups: monopartite replication promoters confined to the leader 3'-end (*Rhabdoviridae* and *Pneumoviridae*) and bipartite replication promoters (*Paramyxoviridae*, *Filoviridae*) harboring a second more internal element separated from the first by the TSS and a spacer sequence of variable length (Fearn and Plemper, 2017; Mühlberger, 2007; Tapparel et al., 1998). Here we studied the length, sequence and structural constraints in the region connecting PE1 and PE2 of the bipartite EBOV promoter, with the primary focus on viral transcription. We could demonstrate, for the first time, that UN<sub>5</sub> hexamer phasing is also essential for Ebola viral transcription, that UN<sub>5</sub> hexamer periodicity can be extended into PE1 and that the spacer region can be expanded by at least 48 nt without losses of transcriptional activity. Making the UN<sub>5</sub> hexamer phasing continuous between PE1 and PE2 enhanced the efficiency of genome transcription and replication.

The EBOV 3'-leader has the potential to form stable secondary structures (Fig. 1B), raising the question whether not only sequence but also structure of PE1 contribute to promoter recognition by the EBOV polymerase. Point mutations in the 3'-terminal hairpin stem (nt -1 to -45; Fig. 1B, genomic RNA) which disrupt single base pairs had no or only a moderate effect on viral polymerase activity; the strongest effect was seen for a G<sub>-9</sub> to C mutation. Likewise, double and triple mutations weakening this hairpin stem showed at best moderate effects on viral transcription/replication (Crary et al., 2003). Employing a minigenome setup similar to the one used in the study presented here, Mühlberger and coworkers (Weik et al., 2005) found that changing four consecutive U residues (-10 to -13, Fig. 1B) to adenosines abolished replication; this also pertained to a construct with four additional compensatory mutations (A to U) restoring base pairing (nt -36 to -39; Fig. 1B) in the stem. Likewise, changing the base identity of nucleotides 44-46, 50-52 or 53-55 abolished the production of replicated RNA (Weik et al., 2005). Together with the results of Crary et al. (2003), these findings led to the conclusion that the 3'-terminal 55 nucleotides (= PE1) of the genomic leader up to the TSS of the first (NP) gene are essential for replication, but it appears that the PE1 sequence rather than its secondary structure is crucial for promoter function.

In the presented study, we observed that deletion of the two 5'-terminal nucleotides in PE1 (nt -54 and -55; construct GP-2 ( $\Delta$ 2 nt PE1), Fig. 5B) immediately preceding the TSS abolished viral transcription, although this construct adhered to hexamer phasing. One possible interpretation is that PE1 has not only replication but also transcription promoter function, similar to what is known for other NNS RNA viruses (reviewed in Noton and Fearn, 2015). In view of the finding by Weik et al. (2005) that changing the base identity of nucleotides 44-46, 50-52 or 53-55 in particular abolishes the production of replicated RNA (Weik et al., 2005), it is not unlikely that sequences in PE1 are recognized by the viral polymerase in its replication and transcription mode. To our knowledge, the structural requirements of a filoviral transcription promoter were not specifically addressed in any previous studies. Yet another interpretation of the  $\Delta$  -54/-55 phenotype could be that the transcription start site must be in the correct hexamer phase relative to the genome 3'-end. This possibility will be tested in future experiments by single nucleotide substitutions at positions -54 and -55. Generally, the features of PE1 that are recognized by the EBOV polymerase in the replication versus transcription mode need to be unraveled. Remarkably, the EBOV transcription promoter appears more similar to promoters of rhabdoviruses, such as VSV, than that of paramyxoviruses with respect to the following aspect: comparable to EBOV, the entire 3'-terminal 50 nt of the VSV genome, including nt -47 to -50 immediately preceding the TSS, harbor signals for transcription (Whelan and Wertz, 1999). In paramyxoviruses, however, a core

transcription promoter could be defined spanning only the 3'-terminal 12 nt of the viral genome (Cowton and Fearn, 2005; Fearn et al., 2002; Tremaglio et al., 2013). Substitution of nt -13 to -36 was neutral and nucleotide identities at positions -37 to -44 immediately preceding the TSS were not essential for transcription, although contributing to transcription activity (McGivern et al., 2005).

Current models of bipartite 3'-leader promoters, developed for paramyxoviruses (Tapparel et al., 1998; Murphy et al., 1998; reviewed in: Le Mercier and Kolakofsky, 2019; Noton and Fearn, 2015), predict that 3'-terminal and internal promoter elements (resembling PE1 and PE2, Fig 1B) need to be juxtaposed on the same vertical face of the nucleocapsid helix for concerted recognition by the viral polymerase. This model is based on the finding that the paramyxoviral nucleocapsid contains 13 NP molecules per helical turn and NP molecules 14-16 (counted from the 3'-end) bind to the three PE2 hexamers (3'-CN<sub>5</sub> in the case of Sendai virus (SeV)). Displacement of the three PE2 hexamers by deletion of 6 nt in the spacer between PE1 and PE2 abolished SeV genome replication (Tapparel et al., 1998). Similar defects were observed for Simian virus 5 (SV5) upon deletion of 6 to 18 nt in the spacer between PE1 and PE2 (Murphy et al., 1998). On the other hand, 6-nt insertions in this region of the SeV promoter were tolerated, but insertion of 12 nt largely ablated promoter activity (Pelet et al., 1996).

For EBOV, relaxed nucleocapsids formed in the presence of NP alone were inferred to have  $26 \pm 2$  NP subunits per helical turn, whereas those assembled in the presence of NP, VP24, VP35 and VP40 are condensed having 13 or 14 per turn (Sugita et al., 2018; Wan et al., 2017). The relaxed nucleocapsids are considered to be present during active transcription and replication (Hoenen et al., 2019), consistent with viral RNA synthesis taking place in the MG system lacking VP24 and VP40. Taking into account that the EBOV NP protein binds 6 nucleotides per NP monomer (Sugita et al., 2018; Wan et al., 2017) as in other NNS systems, the model of concerted polymerase binding to PE1 and PE2 juxtaposed on the same vertical face of the nucleocapsid helix appears questionable for filoviruses: the spacer preceding PE2 can be expanded by up to 48 nt or 8 NP monomers (L+1 variants, Fig. 4A) without apparent losses in viral transcription; in this construct, PE1 and PE2 are expected to be considerably displaced from each other in the nucleocapsid helix, making the concomitant recognition of PE1 and PE2 hard to imagine. Alternatively, the filoviral RNA polymerase might be more flexible than its paramyxoviral counterparts in being able to loop out longer spacer insertions during recognition of PE1 and PE2; or the extended HP structures can form and protrude from the nucleocapsid in the process of promoter recognition to bring PE1 and PE2 closer to each other.

It is interesting to mention that a Gln202 to Ala mutation in the NP protein of the human parainfluenza virus type 2 (hPIV2, paramyxovirus) allowed promoter utilization of the viral polymerase in the absence of the PE2 element which was then largely independent of a genome length divisible by 6. In contrast, promoter utilization was strictly dependent on PE2 and genome hexamer phasing in the presence of wild type NP (Matsumoto et al., 2017, 2018). This observation provides evidence that PE2 is not essential for promoter function but may rather exert a regulatory function. The authors proposed that Gln202 of hPIV2 NP can form a base-specific contact to the 3'-terminal base of a bound RNA hexamer (particularly to the uracil at the genome 3'-end) to control that viral replication/transcription can only take place when the correct hexamer register is ensured. The finding of 3'-UN<sub>5</sub> hexamer periodicity in the EBOV system suggests that such base-specific contacts of NP might also control promoter utilization by the filoviral polymerase.

Making UN<sub>5</sub> hexamer periodicity continuous between nt -51 and -80 by a single point mutation in the context of the native NP construct increased reporter activity as well as the levels of mRNA and cRNA (Fig. 7). This effect was very similar for a largely destabilized NP HP structure and for a variant predicted to be of similar stability as the native NP HP, suggesting that UN<sub>5</sub> hexamer periodicity is functionally more important than the potential to form a stable hairpin structure. It will be interesting to explore if the enhancement of polymerase activity upon completing UN<sub>5</sub> hexamer periodicity between PE1 and PE2 may occur on the level of polymerase initiation frequency, kinetics of elongation or reduced premature termination of polymerization.

It was reported that EBOV mRNAs contain upstream AUG (uAUG) codons that could potentially attenuate translation initiation at the main AUG. This kind of regulation was found to be most pronounced at the L gene with a relatively short 5'-UTR (80 nt), where the upstream ORF (uORF) coding for a 22-aa peptide overlaps with the L ORF (Shabman et al., 2013). The uORF mitigated L translation under normal conditions but elevated L expression under stress conditions. Such translational effects of uORFs could also potentially affect reporter activities measured in the MG system. In our MG, the Rluc gene was preceded by the ~400 nt long 5'-UTR of the NP gene and structural alterations were confined to the very 5'-end of this 5'-UTR. In the case of our L MG constructs (Fig. 3), the uAUG introduced by the L hairpin encodes peptides of 22 aa (L and L+1 [loop] constructs) and 16 aa (L+1 [stem] construct) that are terminated 360-380 nt upstream of the Rluc AUG and thus do not overlap with the main ORF as it is the case in the native L gene (Shabman et al., 2013). We do not see evidence that the uORF in our L constructs might have had a substantial effect on the presented results, considering

that reporter activities of the L+1 [loop] and L+1 [stem] constructs were in the same range as that of the NP (wt) construct (Fig. 4A). As the Rluc gene was preceded by the ~400 nt long 5'-UTR of the NP gene in all of our MG constructs, any uORFs introduced by insertions or deletions in the transcription start region have their stop codons at considerable distance to the AUG of the Rluc gene, making their direct interference with Rluc translation unlikely. Another possibility is that the sequence and structural alterations introduced into the 3'-leader of our mutant minigenomes might have affected mRNA stability (or translation efficiency) rather than viral transcription per se. At present, we cannot exclude that such an effect might have contributed to the phenotypes of some variants such as VP24+2 (stem) and VP24+2 (loop) (Fig. 4A).

We found that hexamer phasing in the EBOV 3'-leader promoter not only applies to replication, but also transcription, and that UN<sub>5</sub> hexamer periodicity can be extended, with interruption, into PE1. So far, we have shown that the distance between PE1 and PE2 can be expanded by up to 48 nt without substantially affecting viral polymerase activity. How can the results be explained? It appears obvious that the hexamer phase is linked to the capacity of each NP protein to bind exactly 6 nt (Sugita et al., 2018; Wan et al., 2017). In addition, the hexamer phasing pertains to transcription and replication, suggesting that a functional aspect common to both processes is affected. We like to put up two models for discussion. In both models we consider that the discrepancy between the hexamer phase in the leader promoter and EBOV genome length being unequal to  $n \times 6$  may imply that those NP molecules binding to the genome ends may either cover less than 6 nt (Le Mercier and Kolakofsky, 2019), or up to a few of the terminal nucleotides remain uncovered by NP. In the first model we assume that there is essentially no replication or transcription activity (not even abortive transcription) of the EBOV polymerase when hexamer phasing in the promoter is violated. Here, the polymerase binds the ~55 3'-terminal nucleotides (PE1) of the genome to position the genome 3'-end in the active site to be able to initiate RNA synthesis (Fig. 8A). Proper positioning of the enzyme on the template strand is only achieved when NP is assembled in the correct register, directed by the UN<sub>5</sub> hexamers, explaining why 6 or 12 nt of the spacer can be deleted or the spacer be expanded by multiples of 6 (Weik et al., 2005; this study). The mechanism how NP monomers are directed into this register during genomic RNA synthesis is yet unclear as NP is thought to bind RNA in a sequence-independent manner (Sugita et al., 2018). A possibility is that the UN<sub>5</sub> hexamers are preferred NP binding sites and consecutive UN<sub>5</sub> hexamers are cooperatively bound by several NP monomers, based on evidence that each NP

monomer may form an H bond to the 3'-nucleotide of the hexamer bound by the neighboring NP molecule in the nucleocapsid (Sugita et al., 2018).

The second model, which is more speculative, may reflect the transition from the initiation to the elongation mode of the polymerase. (Fig. 8B). It assumes that the enzyme starts to polymerize upon positioning of the genome 3'-end in the active site. When the genome 3'-end leaves the template exit channel, the enzyme may hold on to the first UN<sub>5</sub> hexamers (U-51, U-57); as soon as UN<sub>5</sub> hexamers enter the template entry path, they will be bound as well to trigger communication between the entry and exit path; the hexamer periodicity may synchronize re-assembly at the exit site and dis-assembly at the entry site, possibly involving transfer of NP molecules, which are separated from the template RNA at the entry site, to the protruding UN<sub>5</sub> hexamers at the template exit site. It is further conceivable that NP molecules, dis-assembled at the template entry channel, form a transiently RNA-free NP chain that interacts with the polymerase's surface to mediate coordinated translocation of the enzyme in a hexamer phase-dependent manner (Le Mercier and Kolakofsky, 2019). Expansion sequences in the spacer region may loop out as shown in red. It is unclear whether the 3'-terminal RNA hairpin shown in both models forms transiently or at all. Both models shown in Fig. 8 are consistent with the fact that template strand passage through the active site is a commonality of viral replication and transcription, whereas other functional aspects largely differ between the two processes. In contrast to the replication mode, the transcribing polymerase recognizes gene start and end sequences, adds a 5'-cap and monomethylates the mRNA's 5'-end (formation of a cap 1 structure; Fearn and Plemper, 2017), polyadenylates the transcript and releases the newly synthesized mRNA without NP packaging. In contrast to viral mRNAs, replicative cRNA is fully encapsidated by NP during RNA synthesis in NNS viral systems (Ogino and Green, 2019).

## MATERIAL AND METHODS

### *Cell culture and EBOV-specific minigenome systems*

Human embryo kidney cells (HEK293; authenticated as cell line 293 (DMSZ ACC 305)) were cultivated at 37 °C and 5 % CO<sub>2</sub> in a humidified atmosphere in Dulbecco's Modified Eagle Medium (DMEM) containing 50 U/ml penicillin, 50 µg/ml streptomycin, 2 mM L-glutamine and 10% FCS (all from Thermo Fisher Scientific [TFS]). For transfection, 8 x 10<sup>5</sup> HEK293 cells were seeded in each well of a 6-well plate (Greiner) using the aforementioned medium (3 mL/well) and cultivated for 18-24 h until ~60-80 % of cell

confluency was reached. Then transfection of plasmid mixtures was conducted in the same growth medium using the TransIT-LT1 Reagent (Mirus) without media exchange post transfection according to the manufacturer's recommendations. As described previously (Mühlberger et al., 1999), the transfection mixtures included plasmids encoding the EBOV nucleocapsid proteins NP (125 ng), VP35 (125 ng), VP30 (100 ng), L (1000 ng), the respective T7 promoter-driven EBOV-specific minigenome variant coding for Renilla luciferase (250 ng) as well as a plasmid encoding T7 RNA polymerase (250 ng). Additional transfection of pGL4.13 (Promega) encoding a firefly luciferase was performed for normalization of transfection efficiencies (100 ng). For TransIT-LT1:DNA complex formation, 6 µl TransIT-LT1 Reagent were added to a premix of 1950 ng plasmid DNA and 200 µl Opti-MEM I Reduced-Serum Medium (TFS), followed by incubation at room temperature for 15-30 min before drop-wise addition to the HEK293 cells. Replication-competent minigenomes (RC MGs) included the native 3'-leader and 5'-trailer sequences. In contrast, replication-deficient minigenomes (RD MGs) lacked the terminal 55 nt of the antigenomic trailer promoter, thus restricting the viral polymerase to transcription and synthesis of antigenomic copy RNA (cRNA), but excluding the production of new genomic vRNA. As a consequence, measured reporter activities are solely based on viral transcription in RD MG systems. Reporter activities were measured 48 h post transfection (p.t.) in Renilla and Firefly luciferase assays (both purchased from PJK).

### *Plasmids*

Derivatives of plasmid pCAGGS coding for the Zaire EBOV nucleocapsid proteins (VP30, NP, VP35 and L), the T7 RNA polymerase, the EBOV-specific MG (pANDY 3E5E), as well as an EBOV replication-deficient MG, were described earlier (Hoenen et al., 2006, 2010). The *E. coli* DH5 $\alpha$  strain was used for cloning and propagating of plasmids applying standard microbiological procedures.

### *Construction of replication-competent (RC) and replication-deficient (RD) minigenome variants*

All 3'-leader MG variants are based on the negative-sense EBOV-specific MG pANDY 3E5E (Hoenen et al., 2006). The native NP (wt) hairpin, whose 3'-end coincides with the transcription initiation site, was exchanged or mutated in pANDY 3E5E using *DpnI*-based site-directed mutagenesis techniques (see Fig. S1 for details). The used primers are specified in the Supplementary Table S1. Corresponding replication-deficient MGs were constructed by inserting the 3'-leader variants into a pANDY 3E5E derivative plasmid lacking the terminal 55 nucleotides of the trailer replication promoter at the 3' end of the



antigenome (Hoenen et al., 2010) by use of standard restriction enzyme-based approaches or site-directed mutagenesis. Nucleotide numbers refer to the EBOV Mayinga (Zaire, 1976) sequence, GenBank accession no. AF086833. All constructs were verified by DNA sequencing.

#### *Luciferase assay*

For MG activity measurements, cells were lysed in 200 µL 1× Reaction Lysis Buffer (2x Lysis-Juice; PJK). For measurement of Renilla Luciferase activity, the lysates were diluted 1:50 in ddH<sub>2</sub>O in case of replication-competent MG samples. Diluted lysates (10 µL) were mixed with 50 µL of Renilla Luciferase Reagent (Renilla-Juice Fluid, mixed with coelenterazine substrate in reconstruction buffer according to the manufacturer's protocol; PJK). For cells transfected with replication-deficient MGs, 10 µL of undiluted lysate were mixed with 50 µL of Firefly Luciferase Reagent (Beetle-Juice; PJK). Luciferase activities were measured using a Centro LB 960 luminometer (Berthold Technologies). To account for potential differences in transfection efficiency, Renilla luciferase values were normalized to Firefly luciferase values. Results obtained for MGs carrying the native NP (wt) leader were set to 100 %.

#### *RNA extraction and purification for RT-qPCR experiments*

HEK293 cells were transfected with the components of an EBOV-specific minigenome system (Biedenkopf et al., 2013). Cells were harvested at 48 h p.t. and RNA was isolated using the RNeasy mini kit (QIAGEN) according to the manufacturer's protocol, including a first on-column digestion for 1 h using the RNase-Free DNase Set (QIAGEN). RNA was eluted in 40 µL RNase-free ddH<sub>2</sub>O. A second DNase-treatment in solution was performed using Ambion™ DNase I in the presence of 20 U RiboLock RNase Inhibitor (both from TFS) at 37 °C for 1 h in a total reaction volume of 60 µL, followed by purification with Roti®-Phenol/Chloroform/Isoamyl alcohol (Carl Roth). RNA was precipitated by addition of three volumes EtOH containing 0.1 M NaOAc (pH 5). The RNA pellet was washed with 70 % EtOH, air-dried and finally redissolved in RNase-free water.

#### *Reverse transcription*

The strategies for the detection and quantification of the different virus-derived RNAs are schematically illustrated in Fig. S2. 500 ng RNA were used for reverse transcription (RT) of either negative strand RNA [vRNA, primer luc (+) 5'-GGC CTC TTC TTA TTT ATG GCG A-3'] or positive strand RNAs [cRNA/mRNA, primer luc (-) 5'-AGA ACC ATT ACC AGA TTT GCC TGA-3']. Primer sequences for reverse transcription of vRNA and

cRNA/mRNA were adapted from Kruse et al. (2018). In addition, the primer RT\_cRNA 5'-CAG TCC TGC CTT TTC TTT TAA TTT TAT C-3', specific for the cRNA trailer region, was used, and Firefly luciferase mRNA (encoded by the cotransfected pGL4.13 plasmid used for normalization of transfection efficiency) was reverse-transcribed with the Random Hexamer Primer set (TFS). Reactions were conducted with the RevertAid H Minus Reverse Transcriptase (TFS) according to the manufacturer's protocol, including a denaturation step at 65 °C for 5 min as recommended for structured RNAs; cDNAs were diluted 1:10 for quantitative real-time PCR and 2 µL of the diluted cDNA (~5 ng) were used in the respective PCR reaction.

#### *Quantitative real-time PCR*

Quantitative real-time PCR was performed in a total volume of 10 µL on a QuantStudio3™ Real-Time PCR System (TFS) using the PowerUp™ SYBR™ Green Master Mix 2x (TFS). For amplification of vRNA and (+)RNA (cRNA and mRNA), primers luc (+) and luc (-) were used (see above), yielding a PCR product of 112 bp. For specific amplification of cRNA, primers RT\_cRNA (see above) and qPCR\_cRNA (5'- CGG TGA TAG CCT TAA TCT TTG TG-3') were used, yielding a PCR product of 118 bp (see Fig. S2 for primer annealing sites). For amplification of Firefly luciferase mRNA, primers qPCR\_FF\_fwd (5'-CGT GCA AAA GAA GCT ACC G-3') and qPCR\_FF\_rev (5'-GGT GGC AAA TGG GAA GTC AC-3') were combined, yielding a PCR product of 108 bp. PCR conditions were chosen according to the manufacturer's fast cycling mode protocol: Uracil-DNA glycosylase (UDG) activation at 50 °C for 2 min, initial denaturation of cDNA at 95 °C for 2 min, followed by 40 cycles with denaturation at 95 °C for 1 s, annealing and extension at 60 °C for 30 s. RNA levels were quantified using the  $2^{-\Delta\Delta CT}$  values method. Firefly luciferase mRNA was used as reference RNA (internal standard) to calculate  $\Delta C_T$  values.  $2^{-\Delta\Delta CT}$  values were determined for each independent experiment (mostly conducted as technical triplicates) and mean values were derived from 3 or more independent experiments. The mean  $2^{-\Delta CT}$  value determined for the NP (wt) MG was then subtracted from the mean  $2^{-\Delta CT}$  value measured for the mutant MG construct to derive the specific mean  $2^{-\Delta\Delta CT}$  value. To determine mRNA levels,  $2^{-\Delta CT}$  values obtained with the specific cRNA primer set were first subtracted from  $2^{-\Delta CT}$  values obtained with the (+)RNA (cRNA+mRNA) primer set. Note that a  $2^{-\Delta\Delta CT}$  value of 1 would correspond to equal RNA levels for NP (wt) and the mutant MG, whereas values below 1 correspond to higher RNA levels for the NP (wt) MG. Ten independent experiments with three NP (wt) replicates each were analyzed for primer efficiency determination based on fluorescence in exponential phase (Ramakers et al., 2003). The real-time PCR efficiencies were 1.96

for Firefly luciferase mRNA, 1.98 for vRNA and cRNA and 2.00 for (+)RNA (cRNA and mRNA).

#### *Quantification and statistical analysis*

All statistical analyses were performed using GraphPad Prism version 8.1.1 for Windows. Statistical details and definition of parameters can be found in the figure legends. The statistical significance level was chosen as 0.05; calculated p values are indicated in the respective figure legends. Measured values that are normally distributed were analyzed by applying the unpaired *t* test using two-tailed P values. Measured values not normally distributed were analyzed by the nonparametric Mann Whitney test.

### **FUNDING AND ACKNOWLEDGEMENTS**

We like to acknowledge experimental and technical support by Hannes Huber and Sören Seidler (initial cloning experiments) and Astrid Herwig (cell culture). This work was supported by the German Research Foundation (DFG), grant CRC 1021, to S. Becker and R. K. Hartmann.

## REFERENCES

- Biedenkopf, N., Hartlieb, B., Hoenen, T., and Becker, S. (2013). Phosphorylation of Ebola Virus VP30 Influences the Composition of the Viral Nucleocapsid Complex. *J. Biol. Chem.* 288, 11165–11174. doi:10.1074/jbc.m113.461285.
- Biedenkopf, N., Lier, C., and Becker, S. (2016a). Dynamic Phosphorylation of VP30 Is Essential for Ebola Virus Life Cycle. *J. Virol.* 90, 4914–4925. doi:10.1128/jvi.03257-15.
- Biedenkopf, N., Schlereth, J., Grünweller, A., Becker, S., and Hartmann, R. K. (2016b). RNA Binding of Ebola Virus VP30 Is Essential for Activating Viral Transcription. *J. Virol.* 90, 7481–7496. doi:10.1128/jvi.00271-16.
- Blakqori, G., Kochs, G., Haller, O., and Weber, F. (2003). Functional L polymerase of La Crosse virus allows in vivo reconstitution of recombinant nucleocapsids. *J. Gen. Virol.* 84, 1207–1214. doi:10.1099/vir.0.18876-0.
- Bukreyev, A. A., Chandran, K., Dolnik, O., Dye, J. M., Ebihara, H., Leroy, E. M., et al. (2014). Discussions and decisions of the 2012–2014 International Committee on Taxonomy of Viruses (ICTV) Filoviridae Study Group, January 2012–June 2013. *Arch. Virol.* 159, 821–30. doi:10.1007/s00705-013-1846-9.
- Burk, R., Bollinger, L., Johnson, J. C., Wada, J., Radoshitzky, S. R., Palacios, G., et al. (2016). Neglected filoviruses. *FEMS Microbiol. Rev.* 40, 494–519. doi:10.1093/femsre/fuw010.
- Calain, P., Monroe, M. C., and Nichol, S. T. (1999). Ebola Virus Defective Interfering Particles and Persistent Infection. *Virology* 262, 114–128. doi:10.1006/VIRO.1999.9915.
- Calain, P., and Roux, L. (1993). The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *J. Virol.* 67, 4822–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8392616>0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC237869.
- Cowton, V. M., and Fearn, R. (2005). Evidence that the respiratory syncytial virus polymerase is recruited to nucleotides 1 to 11 at the 3' end of the nucleocapsid and can scan to access internal signals. *J. Virol.* 79, 11311–22. doi:10.1128/JVI.79.17.11311-11322.2005.
- Cox, R., Pickar, A., Qiu, S., Tsao, J., Rodenburg, C., Dokland, T., et al. (2014). Structural studies on the authentic mumps virus nucleocapsid showing uncoiling by the phosphoprotein. *Proc. Natl. Acad. Sci.* 111, 15208–15213. doi:10.1073/pnas.1413268111.
- Crary, S. M., Towner, J. S., Honig, J. E., Shoemaker, T. R., and Nichol, S. T. (2003). Analysis of the role of predicted RNA secondary structures in Ebola virus replication. *Virology* 306, 210–218. doi:10.1016/S0042-6822(02)00014-4.
- Deflubé, L. R., Cressey, T. N., Hume, A. J., Olejnik, J., Haddock, E., and Feldmann, F. (2019). Ebolavirus polymerase uses an unconventional genome replication mechanism. 116, 8535–8543. doi:10.1073/pnas.1815745116.
- Fearn, R., Peeples, M. E., and Collins, P. L. (2002). Mapping the Transcription and

- Replication Promoters of Respiratory Syncytial Virus. *J. Virol.* 76, 1663–1672. doi:10.1128/jvi.76.4.1663-1672.2002.
- Fearns, R., and Plemper, R. K. (2017). Polymerases of paramyxoviruses and pneumoviruses. *Virus Res.* 234, 87–102. doi:10.1016/j.virusres.2017.01.008.
- Hoenen, T., Groseth, A., and Feldmann, H. (2019). Therapeutic strategies to target the Ebola virus life cycle. *Nat. Rev. Microbiol.* 17, 593–606. doi:10.1038/s41579-019-0233-2.
- Hoenen, T., Groseth, A., Kolesnikova, L., Theriault, S., Ebihara, H., Hartlieb, B., et al. (2006). Infection of Naïve Target Cells with Virus-Like Particles: Implications for the Function of Ebola Virus VP24. *J. Virol.* 80, 7260–7264. doi:10.1128/JVI.00051-06.
- Hoenen, T., Jung, S., Herwig, A., Groseth, A., and Becker, S. (2010). Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription. *Virology* 403, 56–66. doi:10.1016/j.virol.2010.04.002.
- Kolakofsky, D., Pelet, T., Garcin, D., Hausmann, S., Curran, J., and Roux, L. (1998). Paramyxovirus RNA synthesis and the requirement for hexamer genome length: the rule of six revisited. *J. Virol.* 72, 891–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9444980><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC124558>.
- Kruse, T., Biedenkopf, N., Hertz, E. P. T., Dietzel, E., Stalman, G., López-Méndez, B., et al. (2018). The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. *Mol. Cell* 69, 136–145.e6. doi:10.1016/j.molcel.2017.11.034.
- Kuhn, J. H., Amarasinghe, G. K., Basler, C. F., Bavari, S., Bukreyev, A., Chandran, K., et al. (2019). ICTV virus taxonomy profile: Filoviridae. *J. Gen. Virol.* 100, 911–912. doi:10.1099/jgv.0.001252.
- Le Mercier, P., and Kolakofsky, D. (2019). Bipartite promoters and RNA editing of paramyxoviruses and filoviruses. *RNA* 25, 279–285. doi:10.1261/rna.068825.118.
- Liang, B., Li, Z., Jenni, S., Rahmeh, A. A., Morin, B. M., Grant, T., et al. (2015). Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell*. doi:10.1016/j.cell.2015.06.018.
- Matsumoto, Y., Ohta, K., Kolakofsky, D., and Nishio, M. (2017). A Point Mutation in the RNA-Binding Domain of Human Parainfluenza Virus Type 2 Nucleoprotein Elicits Abnormally Enhanced Polymerase Activity. *J. Virol.* 91, 1–14. doi:10.1128/jvi.02203-16.
- Matsumoto, Y., Ohta, K., Kolakofsky, D., and Nishio, M. (2018). The control of paramyxovirus genome hexamer length and mRNA editing. 461–467. doi:10.1261/rna.065243.117.a.
- McGivern, D. R., Collins, P. L., and Fearns, R. (2005). Identification of Internal Sequences in the 3' Leader Region of Human Respiratory Syncytial Virus That Enhance Transcription and Confer Replication Processivity. *J. Virol.* 79, 2449–2460. doi:10.1128/jvi.79.4.2449-2460.2005.
- Modrof, J., Mühlberger, E., Klenk, H.-D., and Becker, S. (2002). Phosphorylation of VP30 impairs ebola virus transcription. *J. Biol. Chem.* 277, 33099–104.

doi:10.1074/jbc.M203775200.

- Mühlberger, E. (2007). Filovirus replication and transcription. *Future Virol.* 2, 205–215. doi:10.2217/17460794.2.2.205.
- Mühlberger, E., Weik, M., Volchkov, V. E., Klenk, H. D., and Becker, S. (1999). Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J. Virol.* 73, 2333–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9971816> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC104478>.
- Murphy, S. K., Ito, Y., and Parks, G. D. (1998). A functional antigenomic promoter for the paramyxovirus simian virus 5 requires proper spacing between an essential internal segment and the 3' terminus. *J. Virol.* 72, 10–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9420195> [Accessed August 30, 2019].
- Noton, S. L., and Fearn, R. (2015). Initiation and regulation of paramyxovirus transcription and replication. *Virology* 479–480, 545–554. doi:10.1016/j.virol.2015.01.014.
- Ogino, T., and Green, T. J. (2019). RNA Synthesis and Capping by Non-segmented Negative Strand RNA Viral Polymerases: Lessons From a Prototypic Virus. *Front. Microbiol.* 10, 1490. doi:10.3389/fmicb.2019.01490.
- Pelet, T., Delenda, C., Gubbay, O., Garcin, D., and Kolakofsky, D. (1996). Partial characterization of a Sendai virus replication promoter and the rule of six. *Virology* 224, 405–14. doi:10.1006/viro.1996.0547.
- Ramakers, C., Ruijter, J. M., Lekan, Deprez, R. H., and Moorman, A. F. M. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* 339, 62–66. doi:10.1016/S0304-3940(02)01423-4.
- Sanchez, A., Kiley, M. P., Holloway, B. P., and Auperin, D. D. (1993). Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res.* 29, 215–240. doi:10.1016/0168-1702(93)90063-S.
- Schlereth, J., Grünweller, A., Biedenkopf, N., Becker, S., and Hartmann, R. K. (2016). RNA binding specificity of Ebola virus transcription factor VP30. *RNA Biol.* 13, 783–798. doi:10.1080/15476286.2016.1194160.
- Shabman, R. S., Hoenen, T., Groseth, A., Jabado, O., Binning, J. M., Amarasinghe, G. K., et al. (2013). An Upstream Open Reading Frame Modulates Ebola Virus Polymerase Translation and Virus Replication. *PLoS Pathog.* 9. doi:10.1371/journal.ppat.1003147.
- Shu, T., Gan, T., Bai, P., Wang, X., Qian, Q., Zhou, H., et al. (2019). Ebola virus VP35 has novel NTPase and helicase-like activities. *Nucleic Acids Res.* 47, 5837–5851. doi:10.1093/nar/gkz340.
- Sugita, Y., Matsunami, H., Kawaoka, Y., Noda, T., and Wolf, M. (2018). Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex at 3.6 Å resolution. *Nature*. doi:10.1038/s41586-018-0630-0.
- Tapparel, C., Maurice, D., and Roux, L. (1998). The activity of Sendai virus genomic and antigenomic promoters requires a second element past the leader template regions:

- a motif (GNNNNN)<sub>3</sub> is essential for replication. *J. Virol.* 72, 3117–28. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9525637> [Accessed August 28, 2019].
- Tekes, G., Rahmeh, A. A., and Whelan, S. P. J. (2011). A Freeze Frame View of Vesicular Stomatitis Virus Transcription Defines a Minimal Length of RNA for 5' Processing. *PLoS Pathog.* 7, e1002073. doi:10.1371/journal.ppat.1002073.
- Tremaglio, C. Z., Noton, S. L., Deflube, L. R., and Fearn, R. (2013). Respiratory Syncytial Virus Polymerase Can Initiate Transcription from Position 3 of the Leader Promoter. *J. Virol.* 87, 3196–3207. doi:10.1128/JVI.02862-12.
- Wan, W., Kolesnikova, L., Clarke, M., Koehler, A., Noda, T., Becker, S., et al. (2017). Structure and assembly of the Ebola virus nucleocapsid. *Nature* 551, 394–397. doi:10.1038/nature24490.
- Weik, M., Enterlein, S., Schlenz, K., and Muhlberger, E. (2005). The Ebola Virus Genomic Replication Promoter Is Bipartite and Follows the Rule of Six. *J. Virol.* 79, 10660–10671. doi:10.1128/jvi.79.16.10660-10671.2005.
- Weik, M., Modrof, J., Klenk, H.-D., Becker, S., and Muhlberger, E. (2002). Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J. Virol.* 76, 8532–9. doi:10.1128/JVI.76.17.8532.
- Whelan, S. P., and Wertz, G. W. (1999). Regulation of RNA synthesis by the genomic termini of vesicular stomatitis virus: identification of distinct sequences essential for transcription but not replication. *J. Virol.* 73, 297–306. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9847333> [Accessed September 5, 2019].

## FIGURE LEGENDS

**Fig. 1:** Secondary structure formation potential at **(A)** transcription start regions of EBOV genes and **(B)** specifically in the genomic 3'-leader and the complementary antigenomic RNA. **(A)** Genomic sequence elements required for transcription (re)initiation are shown in cyan at the top; transcription is initiated opposite to the 3'-terminal C residue of the TSS. Schematic white boxes at the bottom mark the reading frames for EBOV proteins NP, VP35, VP40, GP, VP30, VP24 and L; light gray boxes indicate 5'- and 3'-UTRs, with dark gray areas depicting the position of the predicted secondary structures illustrated above the genome; p, location of leader and trailer promoters (Calain et al., 1999). **(B)** Validated secondary structures forming in the naked genomic leader RNA (top) and its complementary antigenomic RNA (bottom) (Schlereth et al., 2016; Weik et al., 2002, 2005); in the text and following figures, the genomic RNA is abbreviated as vRNA for viral RNA, and the antigenomic RNA as cRNA for complementary or copy RNA; the numbering of nucleotides in the genomic RNA strand is indicated by a minus sign preceding the nucleotide number. Proposed promoter elements 1 and 2 (PE1, PE2) are shown in green letters and the 3'-U residues of the 8 UN<sub>5</sub> hexamers in PE2 are highlighted in pink in the genomic RNA; orange nucleotides mark the spacer sequence between the transcription start sequence (TSS, in cyan) and PE2. The transcription initiation site on the genomic RNA is marked by the vertical arrow. Nucleotide numbering of the genomic RNA starts at the 3'-terminal nt (position -1) that is complementary to position 1 (5'-terminus) of the antigenome. The 3'-terminal G of the genome is shown as small letter to consider the recent finding that the presence of this nucleotide is not essential as the EBOV RNA polymerase initiates polymerization at the C residue preceding the 3'-terminal G residue (Deflubé et al., 2019).

**Fig. 2:** **(A)** Schematic representation of the replication-competent MG used in the presented study. The sequence complementary to the mRNA encoding the Renilla luciferase reporter mRNA is indicated by the gray box; the red segment corresponds to the hairpin structure shown at the top (TSS plus spacer sequence); white boxes mark sequences antisense (as) to the 5'-UTR of the NP mRNA and antisense to the 3'-UTR of the L mRNA. **(B)** Relative reporter activity of replication-competent (RC) MGs in which the wild type (wt) NP hairpin was replaced with hairpin structures (see Fig. 1A and 3 for details) derived from the transcription start region (TSR) of internal EBOV genes. Data ( $\pm$  standard error of the mean, SEM), normalized to the native NP (wt) construct, are based on at least 3 biological replicates with 2 or 3 technical replicates each. **(C)** Schematic representation of the 3' leader promoter region depicting an extended UN<sub>5</sub> hexamer pattern from PE2 to position -51 in PE1 with a single interruption at G<sub>-75</sub>,



thus defining a reference point (nt -51) in PE1 and linking UN<sub>5</sub> hexamer phasing between PE1 and PE2. Accordingly, the distance between nt -51 and -80 is  $5 \times 6 = 30$  nt in the wt MG (light gray box on the left). Only the mutant construct with the VP40 HP maintained a hexamer phasing (66 nt) between PE1 and PE2, whereas mutant MG constructs harboring the VP35, GP, VP30, VP24 or L HP, respectively, resulted in spacings not divisible by 6 (light gray box on the right). For color coding of the sequence, see legend to Fig. 1B.

**Fig. 3:** Illustration of EBOV 3'-leader MG constructs in which the NP HP (turquoise-shaded area in the center) consisting of the TSS and spacer sequence was replaced with corresponding HP structures derived from the transcription reinitiation sites of internal EBOV genes. For the basic color coding and promoter structure, see legend to Fig. 1B. Wild type MG in the center: U residues that extend the PE2 hexamer phasing 3' of nt -81 are shown in pink; G<sub>-75</sub> interrupting UN<sub>5</sub> hexamer continuity is shown in boldface blue. The peripheral mutant constructs are depicted accordingly, with residues disrupting UN<sub>5</sub> hexamer continuity highlighted in boldface blue as well.

**Fig. 4:** (A) Engineering of the spacer region to conform to hexamer phasing and (B) reverse engineering of the NP (wt) spacer region to violate hexamer phasing in RC MGs. (A) Luciferase reporter activities as a function of sequence/structure variation in the TSS/spacer region of MG 3'-leaders. Leaders containing the NP (wt) HP hairpin or native HP structures of internal EBOV genes instead are indicated by blue bars; HP structures violating hexamer phasing were engineered to restore the hexamer phase by deletion or insertion of nucleotides in the spacer region, either in the apical loop, stem or between the stem and PE2 (gray bars). (B) Luciferase reporter activities of RC MGs carrying the NP (wt) HP or variants with 1-nt deletions or insertions (illustrated in the NP HP structure on the right (+: insertion; Δ: deletion) resulting in deviation from hexamer phasing. Values in (A) and (B) were normalized to the NP (wt) 3'-leader as control (100 %). As a negative control (-L), the plasmid encoding the L gene was omitted in transfection of cells with the set of plasmids encoding EBOV proteins relevant to replication and transcription. Mean values (± SEM) are based on three independent experiments with at least three technical replicates each.

**Fig. 5:** Luciferase reporter activities as a function of compliance with hexamer phasing in the 3'-leader of RD MGs. (A) Performance of MG 3'-leaders in which the NP (wt) HP was replaced with HP structures derived from the TSR of internal EBOV genes, normalized to the MG carrying the NP (wt) 3'-leader; bar color: conforming (blue) or not conforming (gray) to the hexamer phase. (B) MG 3'-leaders with HP structures of internal EBOV

genes, which were mutated by deletion or insertion of nucleotides to conform to hexamer phasing (red bars). **(C)** RD MGs carrying the NP (wt) HP or variants with 1-nt deletions or insertions (for more details, see Fig. 4B). Mean values ( $\pm$  SEM) are based on three independent experiments with at least three technical replicates each. \*\*\* $p < 0.001$  (unpaired  $t$  test).

**Fig. 6:** Impact of hexamer phasing in the 3'-leader on transcription and replication. **(A)** Relative reporter activity of RC MGs with replacements of the NP (wt) HP as indicated (for structural details of leader variants, see Fig. 3 and 4B). **(B)** As in panel A, but the same 3'-leader variants integrated into the RD MG scaffold. Mean values ( $\pm$  SEM), normalized to the native NP (wt) leader construct, are based on 3 biological replicates with 2 or 3 technical replicates each. Above each column, we calculated % activity of the individual mutant MG relative to the NP (wt) MG (= 100 %) after subtraction of the -L background from both. The diagrams in panels A and B combine results already shown in Fig. 4 and 5. **(C)** Mean  $2^{-\Delta C_T}$  values obtained with RNA isolated from cells transfected with the NP (wt) MG to illustrate the abundance of viral mRNA (pink), cRNA (magenta) and vRNA (green), measured by a 2-step strand-specific RT-qPCR of RC MG samples using the same cells as in panel A; - : L omission controls.  $2^{-\Delta C_T}$  values were calculated for each experiment individually; mean  $2^{-\Delta C_T}$  values ( $\pm$  SEM) were derived from 6 to 16 independent experiments with three technical replicates each.  $2^{-\Delta C_T}$  values for mRNA were calculated as follows, based on the mean  $C_T$  ( $\pm$  SEM) values of  $14.52 \pm 0.15$  for mRNA+cRNA,  $20.97 \pm 0.11$  for cRNA and  $23.42 \pm 0.09$  for firefly luciferase mRNA:  $C_T$  (mRNA+cRNA) minus  $C_T$  (firefly luciferase mRNA), i.e.  $14.52 - 23.42 = -8.9$ , corresponding to a  $2^{-\Delta C_T}$  value of 477.7;  $C_T$  (cRNA) minus  $C_T$  (firefly luciferase mRNA), i.e.  $20.97 - 23.42 = -2.45$ , corresponding to a  $2^{-\Delta C_T}$  value of 5.46; subtracting 5.46 from 477.7 then gives the  $2^{-\Delta C_T}$  value of 472.24 for mRNA. Above the cRNA and vRNA columns (+ L), we calculated the mole percentage for cRNA and vRNA relative to mRNA (pink column) after subtraction of the corresponding -L background controls. For the binding sites of primer pairs, see Fig. S2. **(D-F)** Relative levels of viral mRNA, cRNA and vRNA using the same cells as in panel A and normalized to the NP (wt) construct. Mean  $2^{-\Delta \Delta C_T}$  values ( $\pm$  SEM) were derived from at least three independent experiments with two or three technical replicates each. Dotted horizontal lines mark the level of the -L control. For experimental details, see Material and Methods.

**Fig. 7:** UN<sub>5</sub> hexamer continuity between PE1 and PE2 increases viral transcription and replication. **(A)** Sequence and predicted structure of the wt NP HP (left), variant NP U<sub>-75</sub>

and the stabilized variants NP G<sub>-72</sub> and NP U<sub>-75</sub>/G<sub>-72</sub>; the red arrow indicates the mutation that extends UN<sub>5</sub> hexamer phasing and the black one marks the mutation of position -72 that stabilizes the HP stem by converting the A,C mismatch to a G:C bp. The number of UN<sub>5</sub> hexamers between nt -51 and -80 is indicated below each individual structure. The  $\Delta G$  values of the stem-loops are the minimum free energies of the centroid secondary structures predicted by RNAfold for the sequence of nt -52 to -81 using the default parameters (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>). The dashed line connecting A<sub>-65</sub> and U<sub>-69</sub> indicates base pairing as suggested by RNAfold. **(B)** Corresponding luciferase activities of RC (left panel) or RD (right panel) MGs carrying the NP variants illustrated in panel A. Activity values are given in % relative to the native 3'-leader [NP (wt) = 100 %]. As a negative control, the plasmid encoding the L gene was omitted (-L) during transfection. **(C)** Corresponding levels of viral mRNA, cRNA and vRNA measured by a 2-step strand-specific RT-qPCR of RC MG samples using the same cells as in panel B. For more details, see Material and Methods. In panels B and C, mean values ( $\pm$  SEM) were derived from three independent experiments with at least three technical replicates each. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ ; n.s., not significant (unpaired  $t$  test).

**Fig. 8:** Models to explain the hexamer phasing, suggesting a key role of NP which binds 6 nt of RNA per monomer. The sketches are drawn according to the cryo-EM structure of the VSV polymerase (Liang et al., 2015). **(A)** This model assumes that there is essentially no replication or transcription activity (not even abortive transcription) of the EBOV polymerase when hexamer phasing is abolished. Here, the polymerase binds the ~55 3'-terminal nucleotides (PE1) of the genome. Proper positioning of the enzyme on the template strand is only achieved when NP is assembled in the correct register, directed by the UN<sub>5</sub> hexamers, explaining why the spacer can be expanded by multiples of 6. **(B)** This model, which is mechanistically largely speculative, may reflect the transition from the initiation to the elongation mode of the polymerase. It assumes that the enzyme starts to polymerize upon positioning of the genome 3'-end in the active site; when the genome 3'-end leaves the template exit channel, the enzyme may hold on to the first UN<sub>5</sub> hexamers (U-51, U-57); as soon as UN<sub>5</sub> hexamers enter the template entry path, they will be bound as well to trigger communication between the entry and exit path; hexamer periodicity may synchronize re-assembly at the exit site and dis-assembly at the entry site; expansion sequences in the spacer region that may loop out are shown in red. The nascent mRNA transcript is shown in blue; in the VSV system, the nascent mRNA transcript (shown in blue) is capped when reaching a length of 31 nt (Tekes et al., 2011). For more details, see text.

## FIGURES

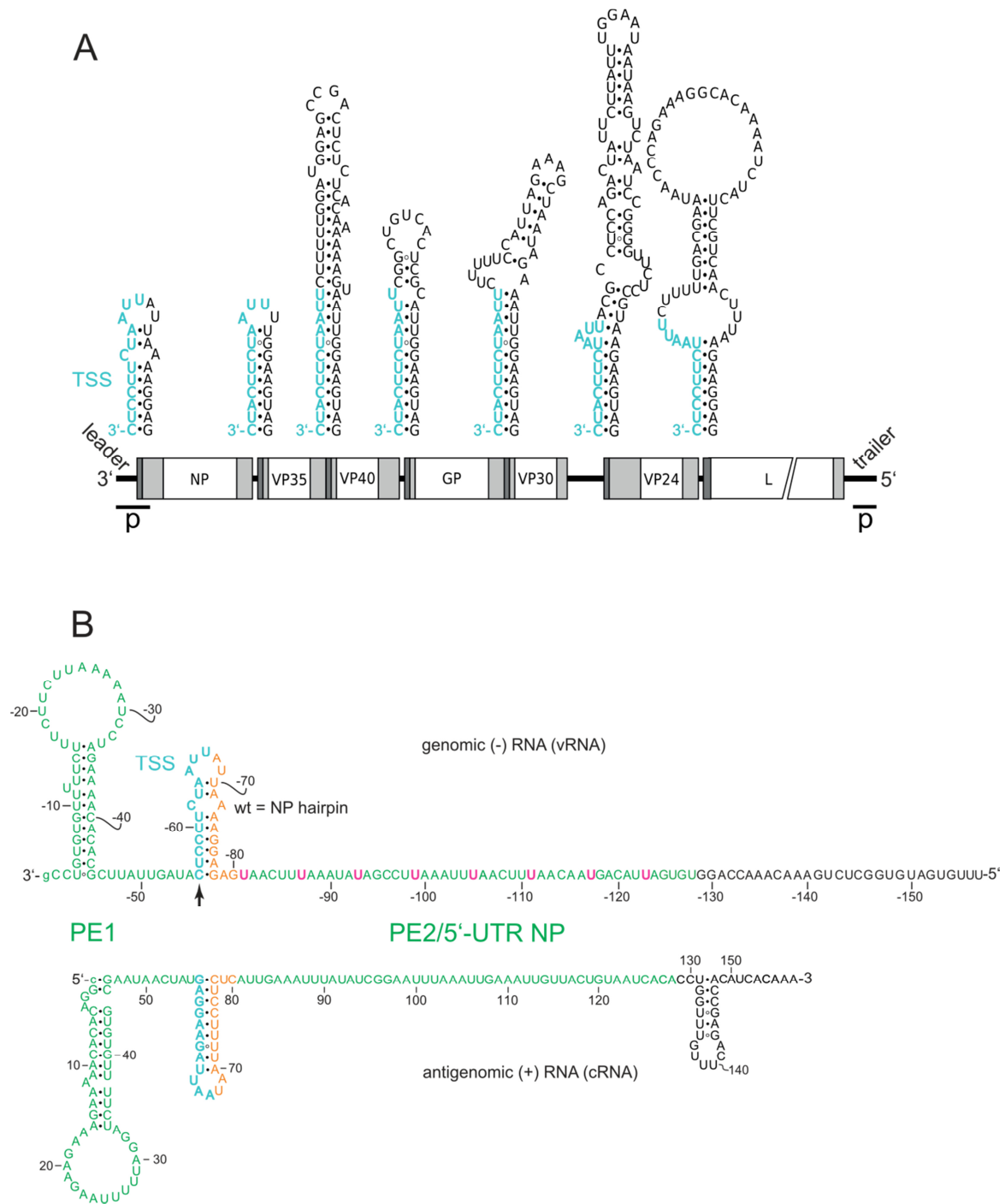


Fig. 1

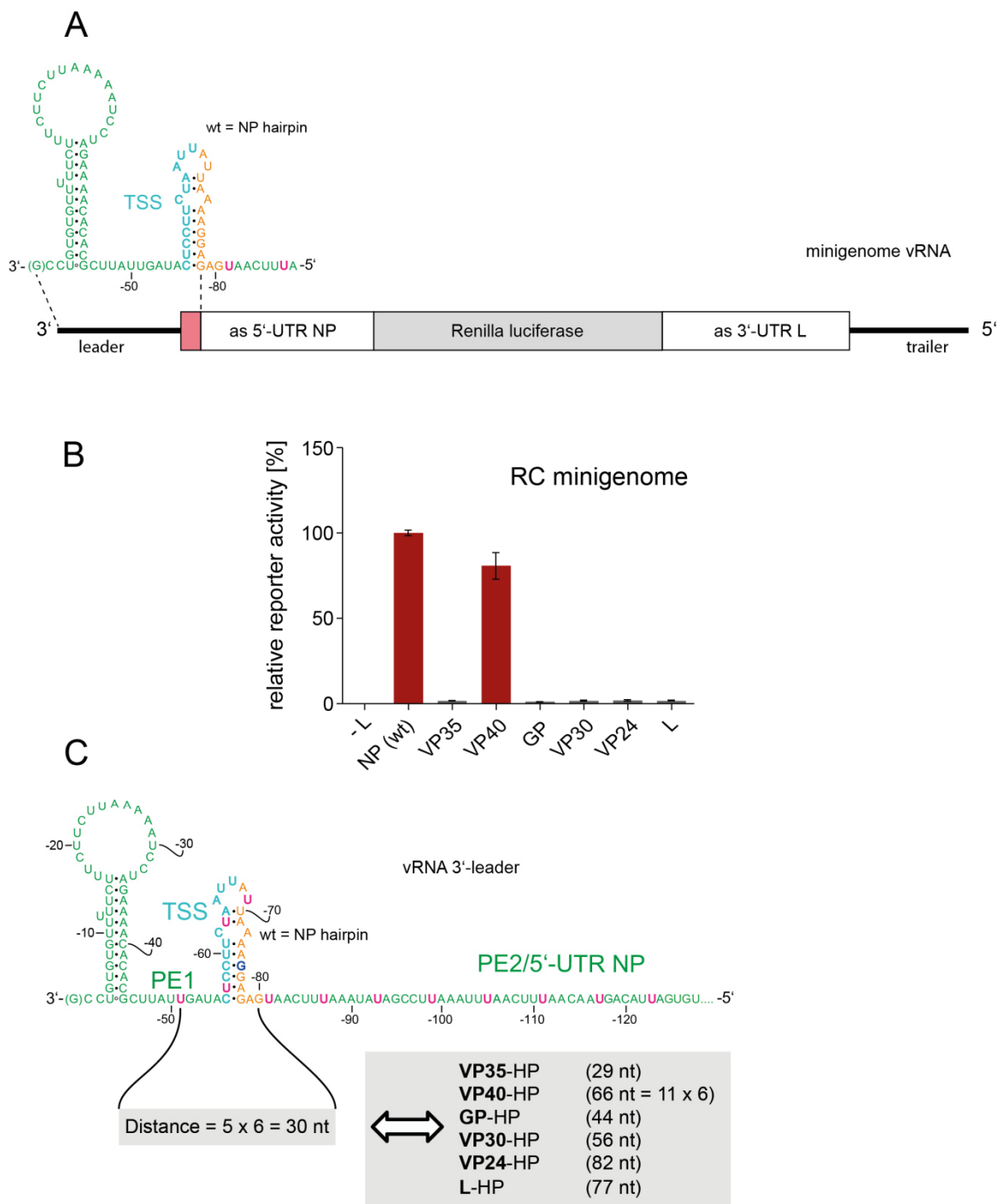


Fig. 2

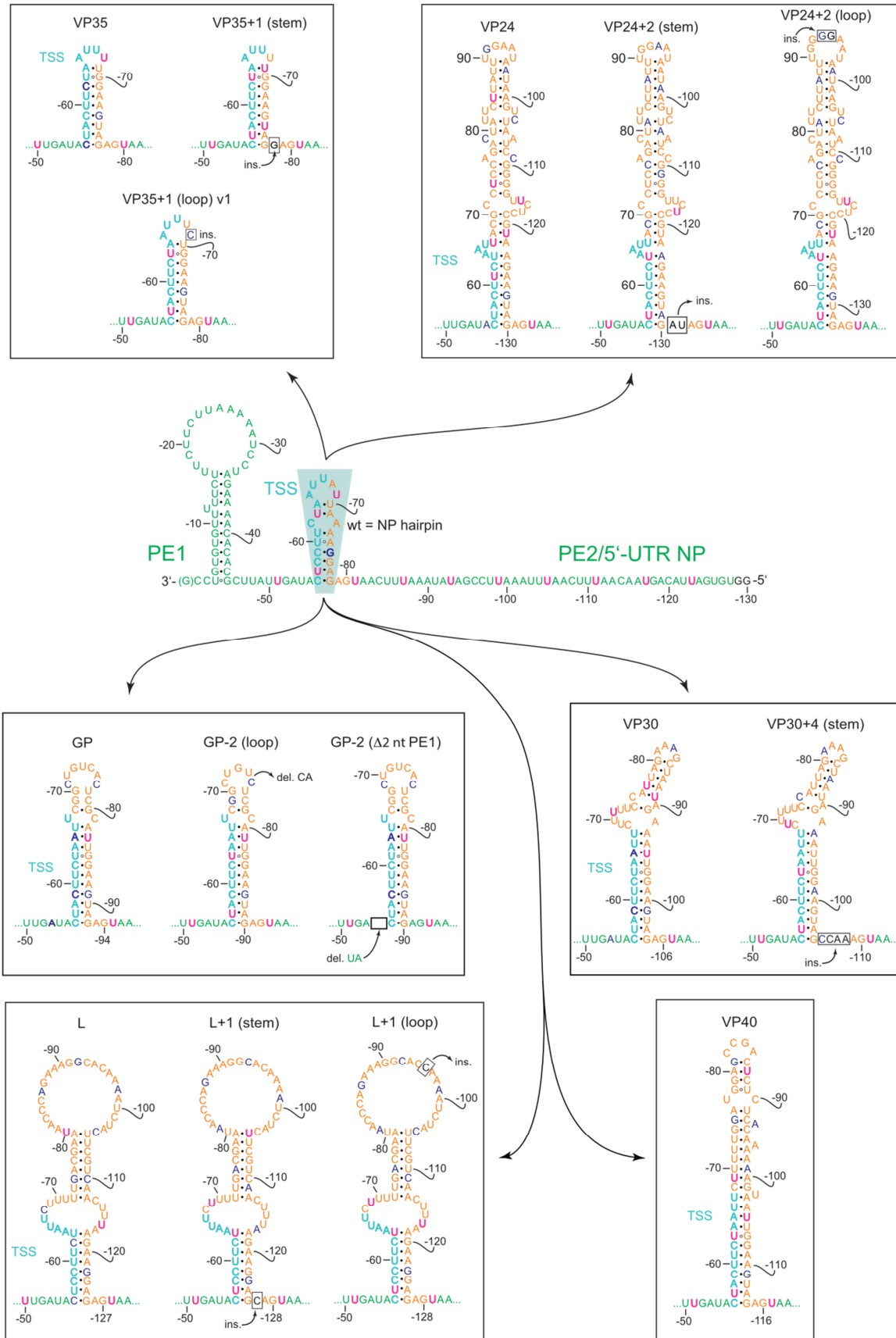


Fig. 3

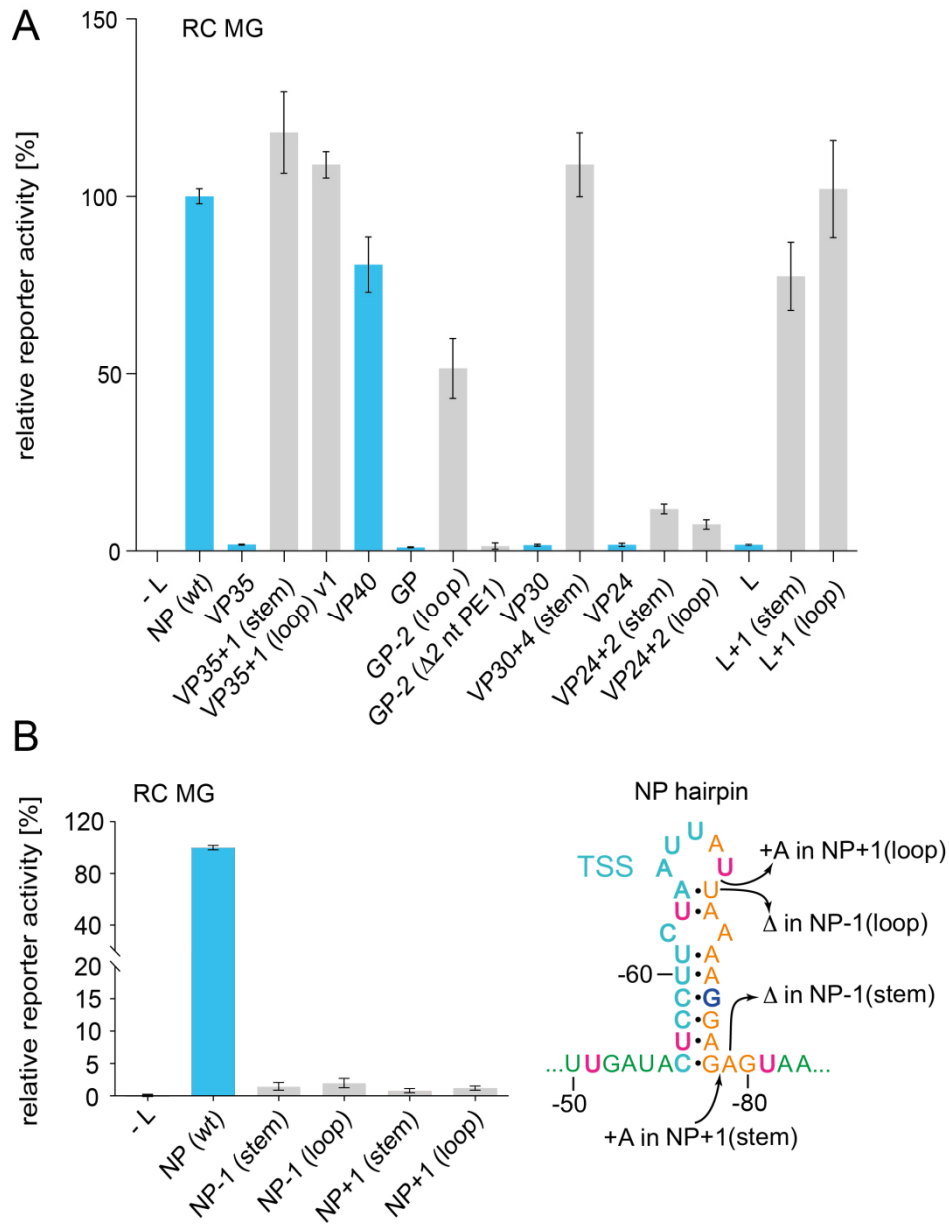


Fig. 4

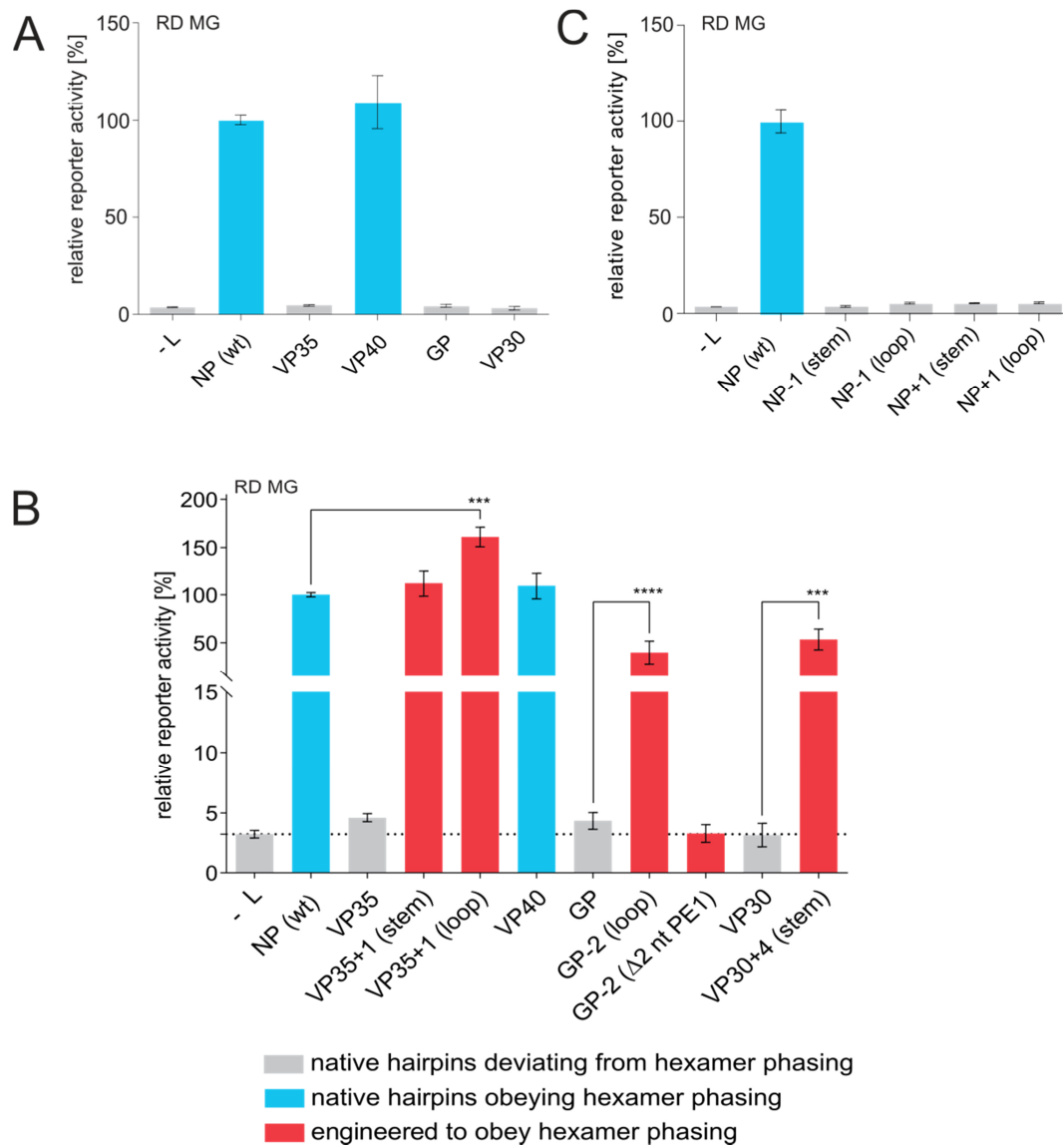


Fig. 5



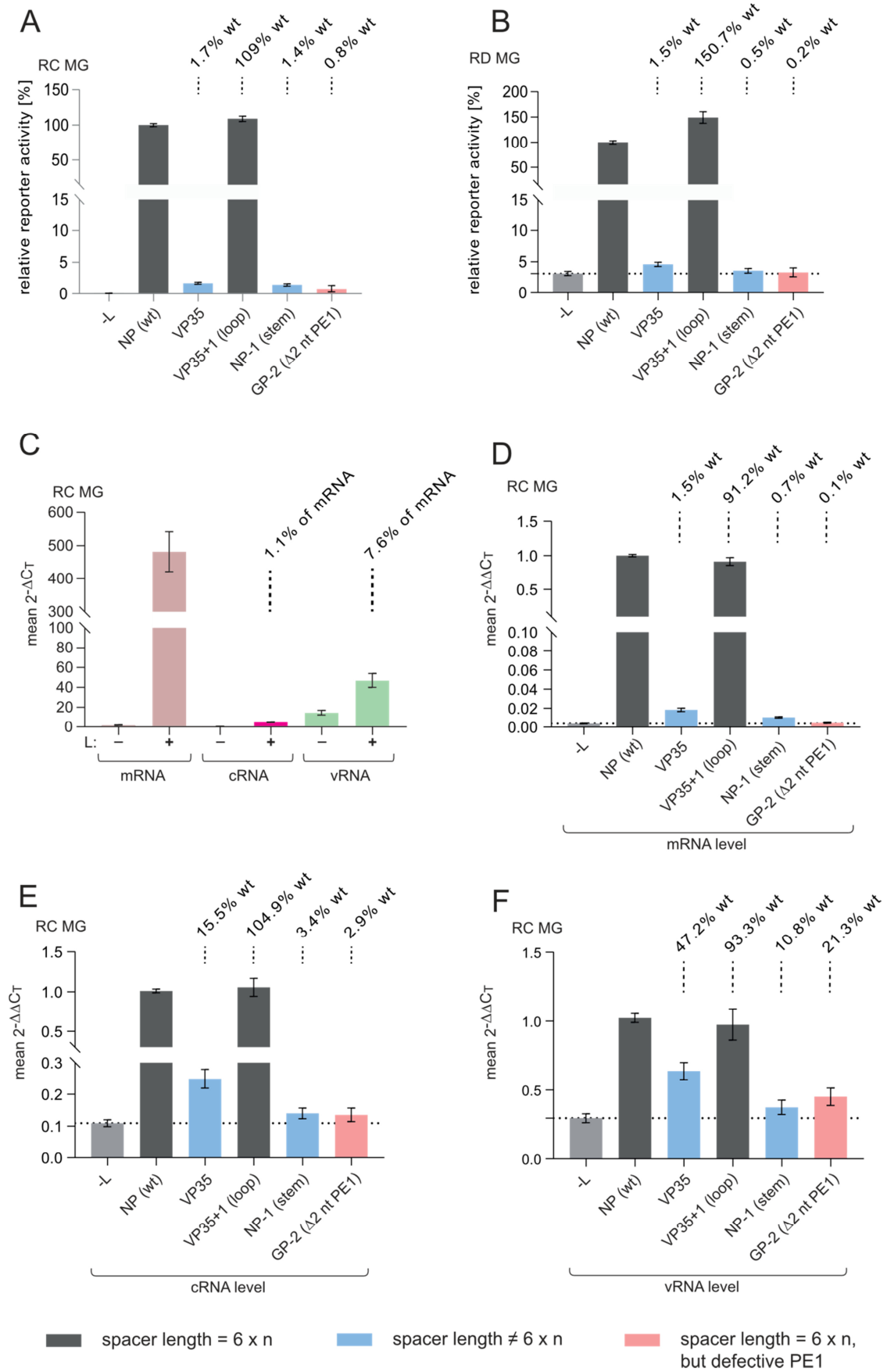
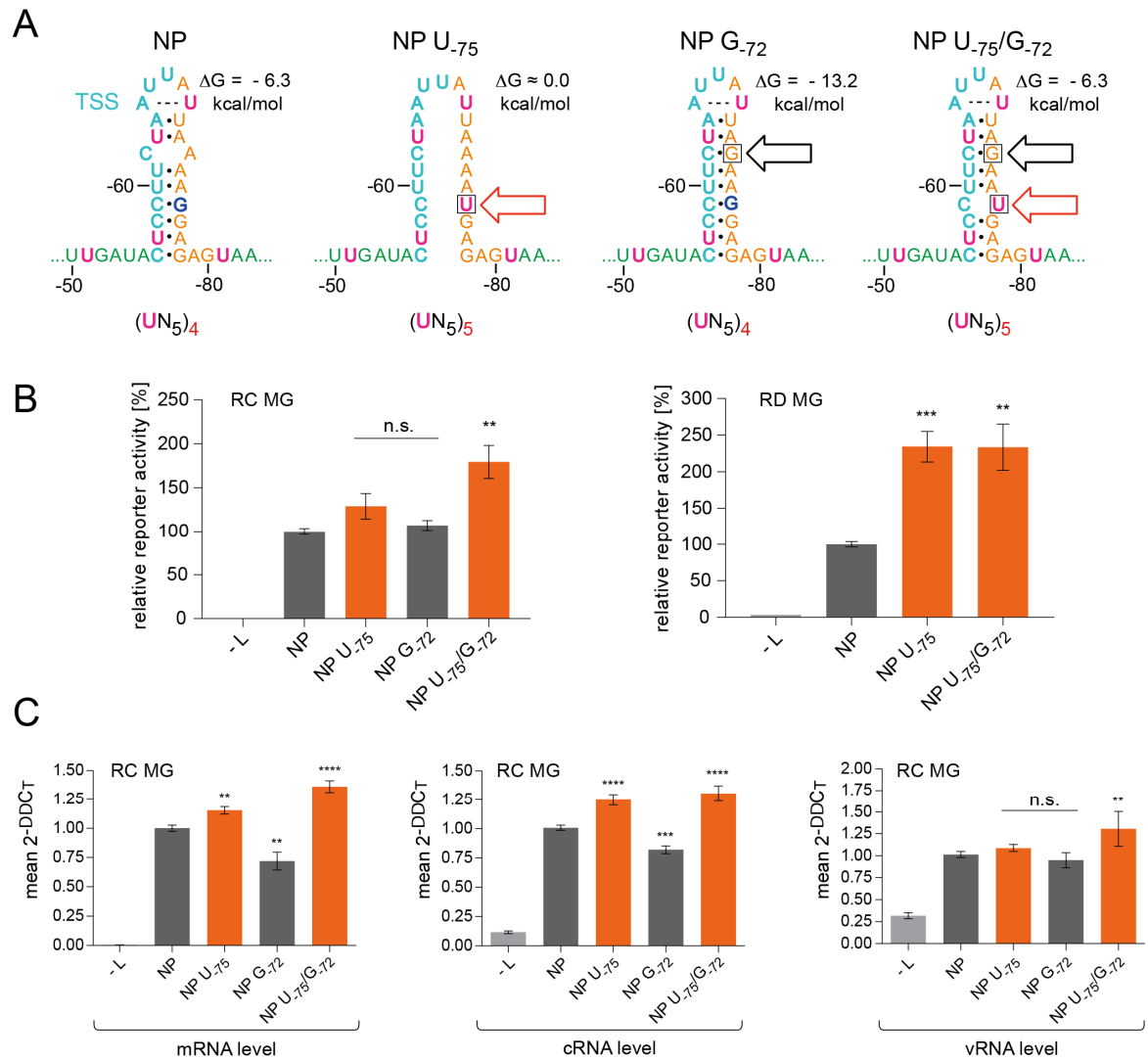


Fig. 6



**Fig. 7**



## **SUPPLEMENTARY MATERIAL**

### **Hexamer phasing governs transcription initiation in the 3'-leader of Ebola virus**

Simone Bach<sup>1</sup>, Nadine Biedenkopf<sup>2</sup>, Arnold Grünweller<sup>1</sup>, Stephan Becker<sup>2</sup>, Roland K. Hartmann<sup>1,#</sup>

<sup>1</sup>Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35037 Marburg, Germany

<sup>2</sup>Institut für Virologie, Philipps-Universität Marburg, Hans-Meerwein-Str. 2, 35043 Marburg

**Tab. S1:** The different PCR-based strategies for site-directed mutagenesis are illustrated in Fig. S2 (except for the classical restriction cloning). The EBOV wt minigenome pANDY 3E5E is termed p3E5E\_Rluc\_RC; RC indicates the replication-competent and RD replication-deficient minigenomes. Hyphens in primer sequences mark nucleotide deletions, boldface underlined nucleotides substitutions or insertions.

**Tab. 36: Cloning strategies and primers of manuscript 1.**

no.	name p3E5E_	Cloning strategy 1:site-directed mutagenesis / plasmid template	primer fwd name	primer forward (5' to 3')	primer rev name	primer reverse (5' to 3')
Overhang/inside-out primer insertion/substitution mutagenesis						
1	VP40	mutagenesis of p3E5E_Rluc_RC	pS_0006	<u>CGAGGTAGGTTTT</u> <u>CTTAATCTTCATCAT</u> AGTTATTCGCACAC AAAAGATCC	pS_000 7	<u>GCTGAGAGAGTGT</u> <u>TCATTAACCTTCATCTC</u> ATTGAAATTTATATCGG AATTTAAATTGAAATTG
2	GP	mutagenesis of p3E5E_Rluc_RC	pS_0016	<u>CACTGTCGGCTTAA</u> <u>TCTTCATCATAGTTA</u> TTCGCACACAAAAG ATCC	pS_001 7	<u>AGCGTAATCTTCATCTC</u> ATTGAAATTTATATCGG AATTTAAATTGAAATTG
3	VP35	mutagenesis of p3E5E_Rluc_RC	pS_0002	TAATCTTCATCATAG TTATTCGCACACAAA AGATCC	pS_000 3	<u>AAACCTTCATCTCATTG</u> AAATTTATATCGGAATTT AAATTGAAATTG
4	VP30	mutagenesis of p3E5E_Rluc_RC	pS_0008	<u>AGATTACCTTTTTCT</u> TAATCTTCATCATAG TTATTCGCACACAAA AGATCC	pS_000 9	<u>TTCGATTATCTTTAATCT</u> <u>TCATCTCATTGAAATTTA</u> TATCGGAATTTAAATTG AAATTG
5	L	mutagenesis of p3E5E_Rluc_RC	pS_0012	<u>CCAATAAGCAGTTTT</u> <u>TCTTAATCTTCCTCA</u> TAGTTATTCGCACAC AAAAGATCC	pS_001 3	<u>GTCTTTCCGTGTTTTAG</u> <u>ATGAAGCAGTTGAAATT</u> <u>CTTCCTCTCATTGAAAT</u> TTATATCGGAATTTAAAT TGAAATTG
6	VP24	mutagenesis of p3E5E_Rluc_RC	pS_0010	<u>GTTTATTCTTATCAG</u> <u>ACCTCCGCATTAAT</u> <u>CTTCATCATAGTTAT</u> TCGCACACAAAAGA TCC	pS_001 1	<u>CTTATTATTCAGATTAG</u> <u>GCCCCAAGAGGCATTG</u> <u>TTCATCTCATTGAAATTT</u> ATATCGGAATTTAAATT GAAATTG
7	VP35+1 nt (stem)	mutagenesis of p3E5E_Rluc_RC	pS_0002	TAATCTTCATCATAG TTATTCGCACACAAA AGATCC	pS_003 7	<u>AAACCTTCATCTCATT</u> GAAATTTATATCGGAAT TTAAATTGAAATTG
Complementary primer mutagenesis						
8	GP -2 (loop)	deletion mutagenesis of p3E5E_GP	pS_0098	GAGATGAAGATTAC GCTC-TGTCGGCTT AATCTTCATCATAG	pS_009 9	CTATGATGAAGATTAAG CCGACA-GAGCGTAATC TTCATCTC
9	GP-2 (Δ2 nt PE1)	deletion mutagenesis of p3E5E_GP	pS_0096	CTGTCGGCTTAATC TTCATCA-GTTATTC GCACACAAAAGATC C	pS_009 7	GGATCTTTTGTGTGCGA ATAAC-TGATGAAGATT AAGCCGAC AG
10	VP24 +2 (stem)	insertion mutagenesis of p3E5E_VP24	pS_0092	CCGATATAAATTTCA ATGATAGATGAAGA ATGCCTCTTGGGGC C	PS_009 3	GGCCCCAAGAGGCATT CTTCATCTATCATTGAA ATTTATATCGG

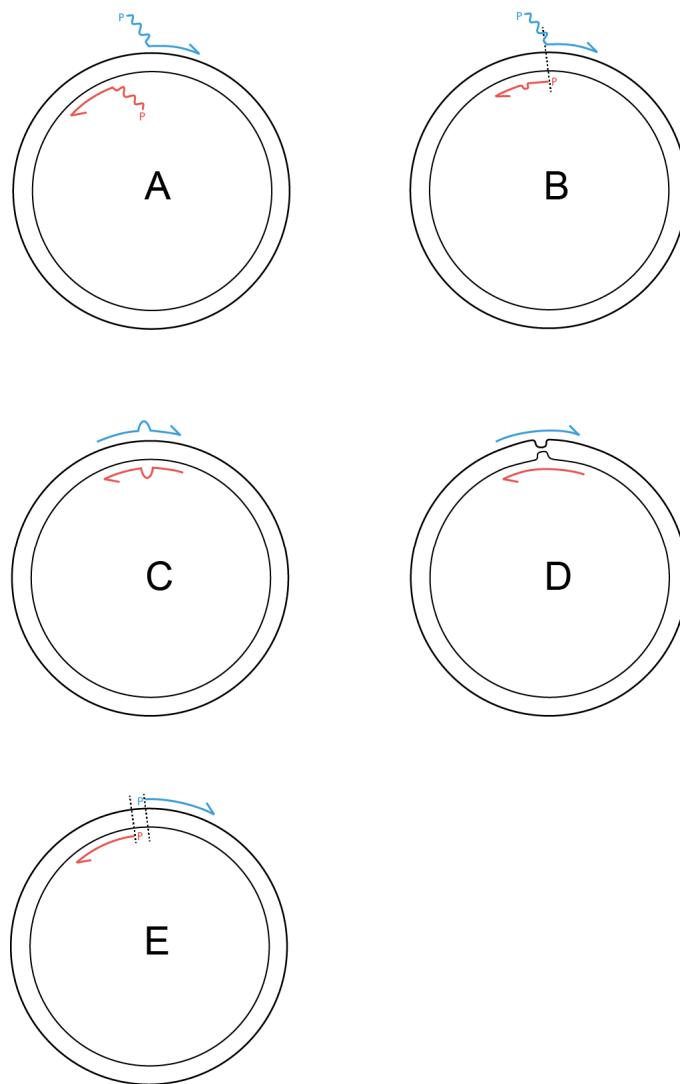
no.	name p3E5E_	Cloning strategy 1:site-directed mutagenesis / plasmid template	primer fwd name	primer forward (5' to 3')	primer rev name	primer reverse (5' to 3')
Complementary primer mutagenesis						
11	VP24 +2 (loop)	insertion mutagenesis of p3E5E_VP24	pS_0094	GGGGCCTAATCTGA ATAATAAGGGGTTTA TTCTTATCAGACCTC CGC	pS_009 5	GCGGAGGTCTGATAAG AATAAACCCCTTATTATT CAGATTAGGCCCC
12	VP35+1 (loop) v1	mutagenesis of p3E5E_VP35	pS_0110	CAATGAGATGAAGG TCTTTAATCTTCATC ATAGTTATTCGC	pS_011 1	GCGAATAACTATGATGA AGATTAAGACCTTCAT CTCATTG
13	VP35+1 (loop) v2	mutagenesis of p3E5E_VP35	pS_227	CAATGAGATGAAGG TTCTTAATCTTCATC ATAGTTATTCGC	pS_228	GCGAATAACTATGATGA AGATTAAGACCTTCAT CTCATTG
14	VP30+4 (stem)	mutagenesis of p3E5E_VP30	pS_0108	CCGATATAAATTTCA ATGAAACCGATGAA GATTAAAGATAATCG	pS_010 9	CGATTATCTTTAATCTTC ATCGGTTTCATTGAAAT TTATATCGG
15	L+1 (stem)	mutagenesis of p3E5E_L	PS_0102	CCGATATAAATTTCA ATGACGAGGAAGAA TTTCAACTGCTTCAT C	PS_010 3	GATGAAGCAGTTGAAAT TCTTCCTCTCATTGAA ATTTATATCGG
16	L+1 (loop)	mutagenesis of p3E5E_L	PS_0104	CAACTGCTTCATCTA AAAACACGGAAGA CCCAATAAGCAG	PS_010 5	CTGCTTATTGGGTCTTT CCGTGGTTTATAGTGAA GCAGTTG
17	VP24+2 (loop) v1	2-step substitution mutagenesis of p3E5E VP24 +2 (loop)	1) pS_214	CTTATCAGACCTCC GTATTAATCTTCATC ATAG	1) pS_215	CTATGATGAAGATTAAT ACGGAGGTCTGATAAG
			2) pS_216	GATATAAATTTCAAT GAGATGAAGAATAC CTCTGGGGCCTAA TC	2) pS_217	GATTAGGCCCAAGAG GTATTCTTCATCTCATT GAAATTTATATC
18	VP24+2 (loop) v2	2-step substitution mutagenesis of p3E5E VP24 +2 (loop)	1) pS_218	GGGGTTTATTCTTAT CAGTCTCCGCATT AATCTTC	1) pS_219	GAAGATTAATGCGGAG GACTGATAAGAATAAAC CCC
			2) pS_220	GCCTCTTGGGGTCT AATCTGAATAATAAG GGG	2) pS_221	CCCCTTATTATTCAGAT TAGACCCCAAGAGGC
19	NP G <sub>72</sub>	substitution mutagenesis of p3E5E_Rluc_RC	pS_229	CAATGAGAGGAAGA TTATTAATCTTCCTC ATAGTTATTCGC	pS_230	GCGAATAACTATGAGGA AGATTAATAATCTTCCT CTCATTG
20	NP-1 (stem)	deletion mutagenesis of p3E5E_Rluc_RC	pS_231	CCGATATAAATTTCA ATG-GAGGAAAATTA TTAATCTTCCTC	pS_232	GAGGAAGATTAATAATT TTCCTC-CATTGAAATT ATATCGG
21	NP-1 (loop)	deletion mutagenesis of p3E5E_Rluc_RC	pS_233	CAATGAGAGGAAAA T-ATTAATCTTCCTC ATAGTTATTCGCAC	pS_234	GTGCGAATAACTATGAG GAAGATTAAT-ATTTCC TCTCATTG
22	NP+1 (stem)	insertion mutagenesis of p3E5E_Rluc_RC	pS_235	CCGATATAAATTTCA ATGAAGAGGAAAAT TATTAATCTTCCTC	pS_236	GAGGAAGATTAATAATT TTCCTCTCATTGAAATT TATATCGG
23	NP+1 (loop)	insertion mutagenesis of p3E5E_Rluc_RC	pS_237	CAATGAGAGGAAAA TATATTAATCTTCCT CATAGTTATTCGCAC	pS_238	GTGCGAATAACTATGAG GAAGATTAATAATTTTC CTCTCATTG

	name p3E5E_	Cloning strategy 1:site-directed mutagenesis / plasmid template	primer fwd name	primer forward (5' to 3')	primer rev name	primer reverse (5' to 3')
Complementary primer mutagenesis						
24	RD_NP-1 (stem)	deletion mutagenesis of p3E5E_Rluc_RD	pS_231	CCGATATAAATTTCA ATG-GAGGAAAATTA TTAATCTTCCTC	pS_232	GAGGAAGATTAATAATT TTCCTC-CATTGAAATTT ATATCGG
25	RD_NP-1 (loop)	deletion mutagenesis of p3E5E_Rluc_RD	pS_233	CAATGAGAGGAAAA T-ATTAATCTTCCTC ATAGTTATTCGCAC	pS_234	GTGCGAATAACTATGAG GAAGATTAAT-ATTTCC TCTCATTG
26	RD_NP+1 (stem)	insertion mutagenesis of p3E5E_Rluc_RD	pS_235	CCGATATAAATTTCA ATGAAGAGGAAAAT TATTAATCTTCCTC	pS_236	GAGGAAGATTAATAATT TTCCTCTTCATTGAAATT TATATCGG
27	RD_NP+1 (loop)	insertion mutagenesis of p3E5E_Rluc_RD	pS_237	CAATGAGAGGAAAA TATTAATCTTCCTC CATAGTTATTCGCAC	pS_238	GTGCGAATAACTATGAG GAAGATTAATAATTTTC CTCTCATTG
28	NP U <sub>-75</sub>	substitution mutagenesis of p3E5E_Rluc_RC	pS_252	CCGATATAAATTTCA ATGAGAGTAAAATTA TTAATCTTCCTCATA G	pS_253	CTATGAGGAAGATTAAT AATTTTACTCTCATTGAA ATTTATATCGG
29	NP U <sub>-75</sub> /G <sub>-72</sub>	substitution mutagenesis of p3E5E_NP G <sub>-72</sub>	pS_250	CCGATATAAATTTCA ATGAGAGTAAAGATT ATTAATCTTCCTCAT AG	pS_251	CTATGAGGAAGATTAAT AATCTTACTCTCATTGA AATTTATATCGG
30	RD_NP U <sub>-75</sub>	substitution mutagenesis of p3E5E_Rluc_RD	pS_252	CCGATATAAATTTCA ATGAGAGTAAAATTA TTAATCTTCCTCATA G	pS_253	CTATGAGGAAGATTAAT AATTTTACTCTCATTGAA ATTTATATCGG
31	RD_NP U <sub>-75</sub> /G <sub>-72</sub>	substitution mutagenesis of p3E5E_Rluc_RD	pS_250	CCGATATAAATTTCA ATGAGAGTAAAGATT ATTAATCTTCCTCAT AG	pS_251	CTATGAGGAAGATTAAT AATCTTACTCTCATTGA AATTTATATCGG
Inside-out primer deletion mutagenesis						
32	RD_VP30	mutagenesis of p3E5E_VP30	pS_0029	TTTCCAGGAATCCTT TTTGCAACG	pS_0030	ACTATAGTGAGTCGTAT TAACCCGGGATCG
33	RD_VP40	mutagenesis of p3E5E_VP40	pS_0029	TTTCCAGGAATCCTT TTTGCAACG	pS_0030	ACTATAGTGAGTCGTAT TAACCCGGGATCG
34	RD_VP35	mutagenesis of p3E5E_VP35	pS_0029	TTTCCAGGAATCCTT TTTGCAACG	pS_0030	ACTATAGTGAGTCGTAT TAACCCGGGATCG
35	RD_GP	mutagenesis of p3E5E_GP	pS_0029	TTTCCAGGAATCCTT TTTGCAACG	pS_0030	ACTATAGTGAGTCGTAT TAACCCGGGATCG
36	RD_VP35+1 (stem)	mutagenesis of p3E5E_VP35+1 (stem)	pS_0029	TTTCCAGGAATCCTT TTTGCAACG	pS_0030	ACTATAGTGAGTCGTAT TAACCCGGGATCG

no.	name	Cloning strategy 2: restriction cloning	primer fwd name	primer fwd (5' to 3')	primer rev name	primer rev (5' to 3')
1) PCR with respective RC plasmid ( no. 8, 9, 12, 14) as template. 2) Restriction cloning with <i>Bam</i> HI and <i>Bst</i> BI into vector p3E5E_Rluc_RD.						
37	RD_GP-2 (loop)		pS_1	CTTTCGAAGTCATG GTGGTATGAGAC	pS_2	GTTAGCAGCCGGATCC TTTTTTTTTG
38	RD_GP-2 ( $\Delta$ 2 nt PE1)		pS_1	CTTTCGAAGTCATG GTGGTATGAGAC	pS_2	GTTAGCAGCCGGATCC TTTTTTTTTG
39	RD_VP35+1 (loop) v1		pS_1	CTTTCGAAGTCATG GTGGTATGAGAC	pS_2	GTTAGCAGCCGGATCC TTTTTTTTTG
40	RD_VP30+4 (stem)		pS_1	CTTTCGAAGTCATG GTGGTATGAGAC	pS_2	GTTAGCAGCCGGATCC TTTTTTTTTG

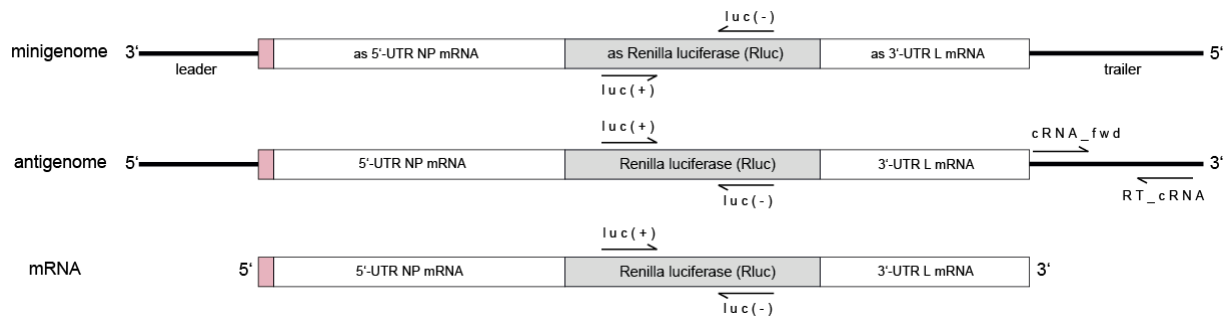


**Fig. S1:** PCR-based strategies (essentially performed as previously described, Li et al., 2009) that were used for the construction of mutant minigenomes. **(A, B)** Overhang/inside-out primer insertion/substitution mutagenesis. **(C, D)** Complementary primer mutagenesis; insertions and substitutions were introduced according to scheme C and deletions according to scheme D. **(E)** Inside-out primer deletion mutagenesis. For more information, see also the NEB web site for the Q5 Site-Directed Mutagenesis Kit (<https://www.neb-online.de/en/pcr-and-dna-amplification/q5-site-directed-mutagenesis-kit/>). In approaches A, B and E, the entire plasmid is amplified with 5'-phosphorylated primers that introduce the desired mutations/substitutions/insertions/deletions, followed by circularization of PCR products and template removal by *DpnI* treatment before bacterial transformation. In approaches C and D, phosphorylation of 5'-ends and ligation are carried out by bacterial enzymes after DNA transformation.



**Fig. S1**

**Fig. S2:** Schematic representation of primers used for qRT-PCR of viral RNA species; as, antisense. See Fig. 2A of the main text for the color code.



**Fig. S2**

## Reference

- Li, D., Willkomm, D. K., and Hartmann, R. K. (2009). Minor changes largely restore catalytic activity of archaeal RNase P RNA from *Methanothermobacter thermoautotrophicus*. *Nucleic Acids Res.* 37, 231–242. doi:10.1093/nar/gkn915.

## 4.2 Einfluss von Sekundärstrukturstabilität und Spacer-Länge auf die EBOV-Transkription und -Replikation

Die Ebola-Virus (EBOV)-Transkription folgt einem Stopp-Start-Mechanismus, bei dem Genstart (GS)- und Genend (GE)-Signale vom viralen Polymerasekomplex erkannt werden. Bemerkenswerterweise haben alle EBOV-GS-Signale das Potenzial, Hairpin-Strukturen sowohl auf genomischer als auch antigenomischer/mRNA-Ebene auszubilden. Die Hairpin-Strukturausbildung am ersten NP-GS-Signal vermittelt die stringente VP30-Abhängigkeit der Transkription. Darüber hinaus ist der NP-Hairpin Teil einer Spacer-Region, die die zwei Elemente des 3'-Leader-Promotors (PE1, PE2) trennt. In der vorliegenden Studie haben wir die maximalen und minimalen Längen- und Strukturstabilitäts-Limits dieses Spacer-Bereichs untersucht. Wir haben dafür eine Reihe von Minigenomen verwendet, in denen der NP-Hairpin entweder mutiert oder durch putative Hairpins der internen Transkriptionsstartregionen ersetzt wurde. Wir konnten zeigen, dass die EBOV-Polymerase noch eine substanzielle Restaktivität von ca. 25 % aufweist, wenn der Spacer um bis zu 66 Nt verlängert wird. Allerdings scheint das Spacer-Längenlimit vom Struktur- und Sequenzkontext abhängig zu sein. Trotz einer Verkürzung der Spacer-Länge durch Deletion von 12 Nukleotiden war nach wie vor eine virale Transkriptions-/Replikationsaktivität von mehr als 50 % messbar, obwohl gleichzeitig jegliche Ausbildung von RNA-Sekundärstruktur unterbunden wurde. Dies deutet darauf hin, dass RNA-Sekundärstrukturen weder für die virale Replikation noch für die Transkription essenziell sind. Eine schrittweise Stabilisierung basengepaarter Bereiche der Hairpin-Strukturen in der Spacer-Region durch strategische Einzelnukleotid-Mutationen, -Deletionen und -Insertionen führte sehr schnell zu Aktivitätsverlusten bis hin zur Elimination jeglicher Polymeraseaktivität. Die Aktivitätsverluste ließen sich besser mit Hairpin-Stabilisierungen auf antigenomischer als auf genomischer Ebene korrelieren. Wir konnten in dieser Studie erstmalig indirekt nachweisen, dass sich die postulierten Haarnadelstrukturen in der Tat in den EBOV-Transkriptionsstartregionen während der viralen RNA-Synthese ausbilden.

## **Secondary structure formation influences EBOV transcription initiation and replication in a length- and stability-dependent manner**

Simone Bach<sup>1</sup>, Nadine Biedenkopf<sup>2</sup>, Stephan Becker<sup>2, #</sup>, Roland K. Hartmann<sup>1, #</sup>

<sup>1</sup>Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35037 Marburg, Germany

<sup>2</sup>Institut für Virologie, Philipps-Universität Marburg, Hans-Meerwein-Str. 2, 35043 Marburg

#corresponding author:

E-mail:       roland.hartmann@staff.uni-marburg.de

---

**ABSTRACT**

Ebola virus (EBOV) transcription follows a stop-start mechanism recognizing gene start and gene end signals (GS, GE). Interestingly, all GS signals are predicted to form hairpin structures on the genomic and antigenomic/mRNA level. Hairpin formation at the first NP GS was suggested to regulate the VP30 dependency of transcription. Furthermore, this NP hairpin is part of a spacer region separating promoter elements 1 and 2 (PE1, PE2) of the EBOV 3'-leader promoter. Here, we explored the limits for insertions and stabilization of RNA structures in this spacer region, using a series of minigenomes in which the NP hairpin region was mutated or replaced with hairpins derived from internal transcription start regions. We demonstrate that the EBOV polymerase retains substantial activity (~ 25 %) in the presence of spacer extensions of up to 66 nt, although depending on the sequence and structure context. A 12-nt deletion in the spacer region, abrogating any potential to form RNA structures, still allowed > 50 % of viral transcription and replication, indicating that RNA structure formation is neither essential for replication nor transcription. Incremental stabilization of hairpin structures in the spacer region resulted in a sharp activity drop when strategic mutations generated continuous helical segments ( $\geq 7$  bp) enriched in G:C base pairs. Activity losses better correlated with hairpin stabilization on the antigenomic than genomic RNA level. Our findings provide first indirect evidence that predicted hairpin structures in EBOV transcription start regions indeed form during viral RNA synthesis.

## INTRODUCTION

Ebola virus (EBOV) is an enveloped, non-segmented negative strand virus (NNS) and belongs to the family of *Filoviridae* in the order *Mononegavirales* (Amarasinghe et al., 2019; Kuhn et al., 2019a). It is the causative agent of Ebola virus disease (EVD) with an overall fatality rate of 80 % in past outbreaks (Malvy et al., 2019). Vaccines protecting against Ebola are under development and have been used to help control the spread of Ebola in the recent outbreak in the Democratic Republic of the Congo (Malvy et al., 2019). However, until now, there is no licensed antiviral treatment available. At the same time there is limited knowledge on the molecular mechanisms of filoviral genome transcription and replication.

The EBOV 19 kb genome comprises seven genes coding for nine mRNAs and proteins (3 GP variants; Mehedi et al., 2011; Fig. 1A). Replication promoters are encoded at the genome's very 3' and 5' ends that are termed 3'-leader and 5'-trailer. The 3'-leader additionally contains signals for transcription initiation, including the transcription start sequence (TSS) of the first nucleoprotein (NP) gene (Mühlberger et al., 1999; Sanchez et al., 1993; Volchkov et al., 1999). A relaxed helical nucleocapsid, composed of the viral genome in complex with the viral proteins NP, VP35 and VP30, is thought to serve as template for viral RNA synthesis. In the helical nucleocapsid, 6 RNA residues are bound per NP molecule (Sugita et al., 2018; Wan et al., 2017). The viral polymerase L requires VP35 as cofactor which has RNA helicase activity (Shu et al., 2019) and may unwind the nucleocapsid to enable L to access the template RNA (Cox et al., 2014). For efficient transcription, EBOV additionally requires VP30 (Mühlberger et al., 1999), a homohexameric phosphoprotein that preferentially binds single-stranded RNA in conjunction with stem-loop structures (Schlereth et al., 2016). Its capability to bind RNA was shown to be a prerequisite for proper transcription activation (Biedenkopf et al., 2016b). However, transcription activation by VP30 is not only dependent on RNA binding but also on a dynamic phosphorylation cycle (Biedenkopf et al., 2016a; Kruse et al., 2018; Lier et al., 2017; Takamatsu et al., 2019).

It is presumed that EBOV transcription follows a stop-start mechanism, recognizing gene start (GS) and gene end (GE) signals at the gene junctions of each EBOV gene and resulting in the synthesis of the individual monocistronic mRNAs. After termination of each transcript at the respective GE signal, the polymerase is thought to either reinitiate mRNA synthesis at the downstream ORF or to dissociate from the template. It is further assumed that the viral polymerase, once dissociated from the template, is only able to reenter the genome at the 3' promoter. Incomplete reinitiation of transcription at internal

genes could explain the lower mRNA levels particularly for the distal VP30, VP24 and L genes, although a so far assumed continuous decline of mRNA levels from the first gene NP to the last gene, as suggested in Fig. 1A, might be questionable based on more recent NGS data (Albariño et al., 2018; Shabman et al., 2014).

The EBOV genome encodes exceptionally long untranslated regions (UTRs (light grey boxes, Fig. 1A) compared to other viruses of the *Mononegavirales* (Feldmann et al., 1992; Sanchez et al., 1993). To date, the UTR sequences flanking the EBOV ORFs and harboring the GS and GE signals are considered to be the main regulators of transcription (Brauburger et al., 2016; Neumann et al., 2009). Yet, potential stem-loop structures are predicted at the beginning of the 5' UTRs of all filoviral mRNAs on the genomic and antigenomic RNA level (Fig. 1A) (Mühlberger et al., 1996; Sanchez et al., 1993). Formation of hairpin (HP) structures in the naked 3'-leader RNA and its antigenomic complement was confirmed by probing experiments (Schlereth et al., 2016; Weik et al., 2002, 2005). However, as the genomic and antigenomic RNA are assumed to be mostly encapsidated by NP, secondary structure formation *in vivo* may only occur on the mRNA level, or may be restricted to short time slots during replication and transcription, when NP proteins have to be removed for threading the template RNA through the active site of the polymerase. The relevance of RNA secondary structure formation to EBOV transcription and replication has not been studied intensively yet. Evidence has been provided that secondary structure formation in the 3'-terminal 45 nt of the genome is not critical for replication (Crary et al., 2003; Weik et al., 2005). However, HP formation at the transcription start region (TSR) of the NP gene was reported to regulate the VP30 dependency of viral transcription. This was inferred from the finding that mutational destabilization of the HP, particularly on the mRNA level, relaxed the VP30 dependency of transcription (Weik et al., 2002).

The NP HP structure separates promoter elements 1 and 2 (PE1 and PE2) of the bipartite EBOV replication promoter (Fig. 1B) (Weik et al., 2005). It was shown that hexamer phasing in this region and the presence of several consecutive 3'-UN<sub>5</sub> hexamers in PE2 are crucial for productive replication, which gave rise to formulation of the "rule of 6" in EBOV replication (Weik et al. 2005). We recently demonstrated that hexamer phasing in the 3'-leader is not only important for viral replication but also essential for viral transcription. In addition, UN<sub>5</sub> hexamer periodicity was shown to extend into PE1, defined by a distance of  $5 \times 6 = 30$  nt between the genomic 3'-leader positions -51 and -80 (Fig. 1B). We demonstrated that the spacer region (orange nt in Fig. 1B) could be extended by up to 48 nt without significant activity losses in transcription and replication of an EBOV minigenome, provided that hexamer

phasing was maintained (Bach et al., 2019). In the present study, we explored the limits of spacer length changes in more detail and further analyzed how stabilization of the potential HP structures at the TSR affects activity of the viral RNA-dependent RNA polymerase (RdRp). We used reporter gene assays as a measure of mRNA synthesis by the EBOV RdRp in combination with qRT-PCR to determine the individual levels of mRNA, genomic RNA (vRNA) and antigenomic (cRNA) for correlating effects on transcription and replication.

## RESULTS

### Reporter gene activity is reduced with longer spacers

Previous studies showed that elongation of the spacer sequence between promoter elements PE1 and PE2 is compatible with promoter function, as long as hexamer phasing is obeyed (Bach et al., 2019, Weik et al., 2005). In an approach where we replaced the NP (wt) HP in the 3'-leader (Fig. 1A) with the HP structures derived from the TSR of internal EBOV genes (see Fig. 2A), we demonstrated that extending the spacer between PE1 and PE2 by 36 (VP40 construct) or 48 nt (L+1 construct) gave rise to reporter activities similar to those of the minigenome carrying the wt 3'-leader (Fig. 2B). However, constructs VP24+2 (stem) and VP24+2 (loop) (Fig. 2A), although adhering to the “rule of 6”, gave rise to very weak reporter activities (Fig. 2B). As the VP24 constructs introduced the longest spacer expansion of all tested variants (+ 54 nt), we considered that this may be attributable to a length limit for such insertions. Therefore, we restored hexamer phasing in the original VP24 construct by deleting 4 or 10 nt (constructs VP24-4 and VP24-10) instead of inserting 2 nt (Fig. 3A). Nucleotide deletions were introduced such that the overall stability of the VP24 HP was changed as little as possible to minimize the potential influence of hairpin stability on polymerase activity. Indeed, constructs VP24-4 and VP24-10 rescued reporter activity (Fig. 3C); reporter activity for the VP24-4 construct, which extends the spacer region by 48 nt, reached ~ 75 % of that of the NP (wt) control, thus displaying very similar activity as the L+1 (stem) construct harboring a 48-nt insertion as well (Fig. 3C). Reducing the spacer insertion to 42 nt (VP24-10) further enhanced reporter activity to a level even extending that of the NP (wt) minigenome (Fig. 3C).

To further scrutinize the influence of sequence and structure context on the length limitation for spacer extensions, we varied spacer length in the context of the VP40 HP. Here, we elongated this HP by inserting 6, 12, 18, 24 or 30 nt into the TSR of the VP40 construct that has a spacer already 36 nt longer than the NP (wt) construct (Fig. 3B). Again, insertions were designed to change the stability of the respective secondary



structures on the vRNA and mRNA level as little as possible (Fig. 3B). In contrast to the constructs based on the VP24 HP, those derived from the VP40 HP construct showed an increased tolerance toward insertions. The VP40+18 construct, inserting 54 extra nucleotides, still gave ~ 60% activity, and even for spacer extensions of 60 and 66 nt (constructs VP40+24, VP40+30); reporter activities were still about ~ 25 % of that of the NP (wt) control (Fig. 3C). We conclude that the spacer length limit depends on the sequence and/or structural context.

### **A minimal spacer shortened by 12 nt results in substantial though somewhat reduced viral RNA levels**

In a reciprocal approach, we introduced a deletion of 12 nt into the NP (wt) spacer, yielding construct  $\Delta 5'$ -spacer (Fig. 4A). This is the largest possible deletion under the premise of preserving the TSS, leaving PE2 intact and maintaining hexamer phasing. A similar deletion of 12 nt was previously shown to be compatible with EBOV MG replication, but effects on viral transcription were not addressed (Weik et al., 2005). The  $\Delta 5'$ -spacer construct is also instrumental to analyzing the absence of any potential to form secondary structures in the TSR, neither on the vRNA nor mRNA level. This minimal spacer mutant construct still gave rise to profound reporter gene activity, amounting to ~ 70 % of that of the NP (wt) MG (Fig. 4B). Analysis of viral RNA levels using a qRT-PCR setup described previously (Bach et al., 2019) revealed reduced levels of mRNA and replicative cRNA. The decrease was more pronounced for mRNA levels (1.75-fold) than cRNA (1.4-fold) and vRNA (1.3-fold) (Fig. 4C-E). In summary, an inter-PE1/PE2 spacing reduced to the TSS and devoid of any potential to form secondary structures still gave rise to substantial reporter activity and substantial viral RNA synthesis, although replication and transcription efficiency of the viral polymerase was decreased relative to the wt construct.

### **Minigenome replication and transcription are highly sensitive to viral 5'-UTR secondary structure stabilizations**

We further explored how stabilization of EBOV HP structures in the promoter spacer may influence activity of the viral polymerase complex. We chose the VP40 HP construct as platform for stabilizing mutations, as it represents the most stable authentic EBOV TSR HP on the vRNA level, while at the same time showing only moderate stability on the mRNA level compared with the other potential EBOV 5'-UTR structures (Fig. 1A). We stepwise increased the amounts of G:C and/or A:U base pairs in the hairpin's middle and upper section (elements highlighted in ochre in Fig. 5) while maintaining hexamer phasing and rendering the VP40 TSS and the overlapping VP35 GE signal (magenta line

in Fig. 5, VP40 HP) unchanged. Minigenomes harboring these stabilized VP40 hairpin derivatives (termed VP40<sub>1 to 6</sub>) were then analyzed for reporter activity as well as mRNA, cRNA and vRNA levels (Fig. 6). For comparison of stabilization effects in different sequence and length contexts, we further included a construct with a stabilized NP G<sub>-72</sub> HP (Fig. 5) that was described previously (Bach et al., 2019). Here, a G:C base pair was introduced into the NP HP, resulting in a ~ 2.2-fold and 1.3-fold HP stabilization on the vRNA and mRNA level, respectively (Fig. 5). The NP G<sub>-72</sub> HP construct showed wt-like reporter activity and slightly reduced mRNA and cRNA levels (Fig. 6A-C). For the stabilized constructs VP40<sub>6</sub> and VP40<sub>5</sub>, we observed ~ 50 to 60 % reporter activity compared with ~ 80 % activity for the original VP40 construct (Fig. 6A). However, mRNA levels in particular as well as cRNA levels were already substantially reduced for construct VP40<sub>5</sub> relative to the NP (wt) construct and the original VP40 construct (Fig. 6B, C). Remarkably, a sharp drop almost to background levels was observed for reporter activity and mRNA levels of construct VP40<sub>4</sub> (Fig. 6A, B). A corresponding decrease in cRNA levels was observed as well (cf. Fig. 6B and C). For the most stable variants, VP40<sub>2</sub> and VP40<sub>1</sub>, viral RNA levels completely dropped to background levels (Fig. 6).

## DISCUSSION

For some positive-sense RNA viruses such as the hepatitis C virus or the family of coronaviruses, secondary structures, especially at the genomic 3' and 5' ends, were shown to be critical regulatory elements influencing viral replication and transcription (Lytle et al., 2002; Yang and Leibowitz, 2015). So far there is little evidence for RNA secondary structure formation influencing replication or transcription of NNS viruses. For some NNS viruses, panhandle structures resulting from base pairing of the 3' and 5' terminal regions (similar to those formed in genomic segments of influenza A virus) were shown to be relevant for interaction with the viral polymerase (Barr and Wertz, 2004; Flick and Hobom, 1999; Fodor et al., 1994; Pigott, 2005; Tomescu et al., 2014). On the contrary, for other NNS viruses, including EBOV, potential RNA secondary structure formation at the genomic 3' end or panhandle formation owing to pairing of 3'- and 5'-terminal sequences is either unlikely or not relevant to virus replication (Crary et al., 2003; Hoffman and Banerjee, 2000a; Weik et al., 2005).

First evidence for a role of RNA structure formation in EBOV transcription came from a study in which the potential NP HP at the first TSS was weakened, primarily on the mRNA level, by introducing six point mutations (generating a *NheI* restriction site) but without changing the spacer length between PE1 and PE2. This resulted in a relaxed

VP30 dependency of viral transcription, associated with a trend toward somewhat increased mRNA levels as suggested by Northern blot results (Weik et al., 2002). In the present study, we constructed the  $\Delta 5'$ -spacer mutant that abolished any potential to form RNA structures at the TSR on the vRNA and mRNA level. Active transcription and replication of the  $\Delta 5'$ -spacer construct indicates that RNA secondary structure formation in the TSR is not obligatory for initiation of RNA synthesis by the EBOV polymerase complex. However, reporter activity as well as cRNA, vRNA and particularly mRNA levels were moderately reduced for this construct relative to the NP (wt) construct (Fig. 4). This may be taken as indication that HP formation at the TSS (either on the genomic and/or antigenomic RNA level) contributes to the overall efficiency of RNA synthesis. Alternatively, the altered spatial relationship of PE1 and PE2 as a consequence of the 12-nt deletion may be the primary reason for compromised promoter function.

In the present study, we were able to show that both length and stability of potential HP structures in the PE1-PE2 spacer region of the EBOV 3'-leader promoter influence transcription and replication. The TSR of the VP24 gene encodes the longest HP structure among all EBOV genes, with by far the highest predicted overall stability of  $\Delta G \approx -28$  kcal/mol on the mRNA/cRNA level (Fig. 1A). The same pertains to the VP24+2 variants (Fig. 3A) that were constructed to maintain UN<sub>5</sub> hexamer phasing (Bach et al., 2019). By deleting 4 and 10 nt in the original VP24 construct, yielding constructs VP24-4 and VP24-10 that maintain hexamer phasing, we could increasingly restore promoter function (Fig. 3A, C). For the VP24-4 construct, overall HP stability on the mRNA/cRNA level decreased from -28 to -24 kcal/mol while overall HP stability on the vRNA level was hardly changed (even slightly increased relative to all tested VP24-derived constructs (Fig. 1A, 3A). These observations may suggest that the overall HP stability on the mRNA/cRNA level inversely correlates with RNA synthesis efficiency. This is in line with our finding that longer spacer extensions were tolerated in the context of the VP40 HP compared with the VP24 HP (Fig. 3C), as all VP40 constructs had predicted overall stabilities with  $\Delta G$  values weaker than -23 kcal/mol on the level of both RNA strands (Fig. 3B). In this respect, the VP40 HP turned out to be a more suitable framework toward exploring the length limitation for insertions into the spacer region of the EBOV leader promoter.

Regarding the stabilized VP40 HP-derived constructs, promoter function began to substantially impair when going from variant VP40<sub>6</sub> to VP40<sub>5</sub> and transcription collapsed with variant VP40<sub>4</sub> (Fig. 5 and 6). For constructs 6 vs. 5 vs. 4, HP stabilities changed from -26 vs. -29.9 vs. -29 kcal/mol on the vRNA level and from -19.7 vs. -25.3 and -26.6

kcal/mol on the mRNA/cRNA level (Fig. 5). The activity differences do not correlate well with increases in HP stability on the vRNA level, as variant 4 is predicted to form a slightly less stable HP than variant 5 on the vRNA level, but is much less active than variant 5. Rather, activity losses better correlate with HP stabilization on the mRNA/cRNA level. Thus, there is multiple evidence that spacer sequences, which have the potential to form structures on the mRNA/cRNA level that exceed a certain stability threshold, negatively affect EBOV transcription and replication. A detailed comparison of the more local duplex stabilizations in the various VP40 HP constructs may further sharpen this conclusion. Regarding the steep activity drop of construct VP40<sub>4</sub> relative to VP40<sub>6</sub>, it is remarkable that we solely converted a C-A mismatch (on the mRNA/cRNA level) in VP40<sub>6</sub> to a G:C bp in variant VP40<sub>4</sub>. This mutation generated a continuous helical stretch of 7 bp including 4 G:C bp (Fig. 5, highlighted in ocher) in the apical part of the HP structure. It seems that such continuous helical segments enriched in G:C base pairs and with low positional entropy (Fig. S2) are detrimental to viral RNA synthesis. This opens perspectives to inhibit EBOV transcription and replication by antisense technologies using locked nucleic acids that form particularly stable duplexes with RNA molecules (Grünweller and Hartmann, 2007).

Among the VP40 constructs, variant VP40<sub>5</sub> was the only one that had an additional UN<sub>5</sub> hexamer. Extended UN<sub>5</sub> hexamer phasing in the TSR was recently shown to enhance viral replication and transcription, although in a context-dependent manner (Bach et al., 2019). Thus, we cannot exclude that this feature of construct VP40<sub>5</sub> slightly compensated the negative effect of its HP stabilization on reporter activity and viral RNA levels.

Reporter gene activity measured in replication-competent minigenome systems is determined by viral transcription/replication and translation of viral mRNAs by cellular ribosomes. Thus, low reporter activity seen for construct VP40<sub>6</sub> (Fig. 6A) could be potentially due to inefficient translation of luciferase mRNA carrying a stabilized 5'-UTR. Yet, the qRT-PCR data demonstrate that the primary defect of this variant is on the level of viral transcription, with an additional contribution from reduced cRNA synthesis (Fig. 5B). Our findings provide first indirect evidence that predicted hairpin structures do form during viral RNA synthesis. One possibility is that VP35 via its recently identified RNA helicase activity (Shu et al., 2019) unwinds such 5'-UTR structures when they form cotranscriptionally. However, the helicase may become inefficient on extended and rigidified duplex structures.

## **MATERIAL AND METHODS**

Cell culture, EBOV-specific minigenome assays, Renilla and Firefly luciferase activity assays were carried out exactly as described (Bach et al., 2019). The same pertains to preparation of total RNA extracted from cells transfected with plasmids of the EBOV minigenome system, reverse transcription and quantitative real-time PCR of EBOV RNAs.

### *Quantification and statistical analysis*

All quantifications and statistical analyses were performed using GraphPad Prism version 8.1.1 for Windows, with statistical details and definition of parameters including P values defined in the figure legends. Normally distributed samples were analyzed by applying the unpaired *t* test using two-tailed P values. The nonparametric Mann Whitney test was applied to samples deviating from normal distribution.

## **FUNDING**

This work was supported by the German Research Foundation (DFG), grant CRC 1021, to S. Becker and R. K. Hartmann.

## **ACKNOWLEDGEMENTS**

We like to acknowledge experimental and technical support by Hannes Huber, Astrid Herwig (cell culture) and Arnold Grünweller for fruitful discussions.

## REFERENCES

- Albariño, C. G., Guerrero, L. W., Chakrabarti, A. K., and Nichol, S. T. (2018). Transcriptional analysis of viral mRNAs reveals common transcription patterns in cells infected by five different filoviruses. *PLoS One* 13, 1–13. doi:10.1371/journal.pone.0201827.
- Barr, J. N., and Wertz, G. W. (2004). Bunyamwera Bunyavirus RNA Synthesis Requires Cooperation of 3'- and 5'-Terminal Sequences. *J. Virol.* 78, 1129–1138. doi:10.1128/jvi.78.3.1129-1138.2004.
- Biedenkopf, N., Lier, C., and Becker, S. (2016a). Dynamic Phosphorylation of VP30 Is Essential for Ebola Virus Life Cycle. *J. Virol.* 90, 4914–4925. doi:10.1128/jvi.03257-15.
- Biedenkopf, N., Schlereth, J., Grünweller, A., Becker, S., and Hartmann, R. K. (2016b). RNA Binding of Ebola Virus VP30 Is Essential for Activating Viral Transcription. *J. Virol.* 90, 7481–7496. doi:10.1128/jvi.00271-16.
- Brauburger, K., Boehmann, Y., Kräling, V., and Mühlberger, E. (2016). Structure and Regulatory Elements on Gene Expression and Polymerase Scanning Behavior. *J. Virol.* 90, 1898–1909. doi:10.1128/JVI.02341-15.Editor.
- Bukreyev, A. A., Chandran, K., Dolnik, O., Dye, J. M., Ebihara, H., Leroy, E. M., et al. (2014). Discussions and decisions of the 2012–2014 International Committee on Taxonomy of Viruses (ICTV) Filoviridae Study Group, January 2012–June 2013. *Arch. Virol.* 159, 821–30. doi:10.1007/s00705-013-1846-9.
- Calain, P., Monroe, M. C., and Nichol, S. T. (1999). Ebola Virus Defective Interfering Particles and Persistent Infection. *Virology* 262, 114–128. doi:10.1006/VIRO.1999.9915.
- Cox, R., Pickar, A., Qiu, S., Tsao, J., Rodenburg, C., Dokland, T., et al. (2014). Structural studies on the authentic mumps virus nucleocapsid showing uncoiling by the phosphoprotein. *Proc. Natl. Acad. Sci.* 111, 15208–15213. doi:10.1073/pnas.1413268111.
- Crary, S. M., Towner, J. S., Honig, J. E., Shoemaker, T. R., and Nichol, S. T. (2003). Analysis of the role of predicted RNA secondary structures in Ebola virus replication. doi:10.1016/S0042-6822(02)00014-4.
- Deflubé, L. R., Cressey, T. N., Hume, A. J., Olejnik, J., Haddock, E., and Feldmann, F. (2019). Ebolavirus polymerase uses an unconventional genome replication mechanism. 116, 8535–8543. doi:10.1073/pnas.1815745116.
- Feldmann, H., Mühlberger, E., Randolph, A., Will, C., Kiley, M. P., Sanchez, A., et al. (1992). Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus Res.* 24, 1–19. doi:10.1016/0168-1702(92)90027-7.
- Flick, R., and Hobom, G. (1999). Interaction of influenza virus polymerase with viral RNA in the “corkscrew” conformation. *J. Gen. Virol.* doi:10.1099/0022-1317-80-10-2565.
- Fodor, E., Pritlove, D. C., and Brownlee, G. G. (1994). The influenza virus panhandle is involved in the initiation of transcription. *J. Virol.* 68, 4092–6. Available at:

- <http://www.ncbi.nlm.nih.gov/pubmed/8189550>.
- Grünweller, A., and Hartmann, R. K. (2007). Locked Nucleic Acid Oligonucleotides. *BioDrugs* 21, 235–243. doi:10.2165/00063030-200721040-00004.
- Hoffman, M. A., and Banerjee, A. K. (2000). Analysis of RNA secondary structure in replication of human parainfluenza virus type 3. *Virology*. doi:10.1006/viro.2000.0369.
- Kruse, T., Biedenkopf, N., Hertz, E. P. T., Dietzel, E., Stalman, G., López-Méndez, B., et al. (2018). The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. *Mol. Cell* 69, 136–145.e6. doi:10.1016/j.molcel.2017.11.034.
- Li, D., Willkomm, D. K., and Hartmann, R. K. (2009). Minor changes largely restore catalytic activity of archaeal RNase P RNA from *Methanothermobacter thermoautotrophicus*. *Nucleic Acids Res.* 37, 231–242. doi:10.1093/nar/gkn915.
- Lier, C., Becker, S., and Biedenkopf, N. (2017). Dynamic phosphorylation of Ebola virus VP30 in NP-induced inclusion bodies. *Virology* 512, 39–47. doi:10.1016/J.VIROL.2017.09.006.
- Lytle, J. R., Wu, L., and Robertson, H. D. (2002). The Ribosome Binding Site of Hepatitis C Virus mRNA. *J. Virol.* 75, 7629–7636. doi:10.1128/jvi.75.16.7629-7636.2001.
- Malvy, D., McElroy, A. K., de Clerck, H., Günther, S., and van Griensven, J. (2019). Ebola virus disease. *Lancet* 393, 936–948. doi:10.1016/S0140-6736(18)33132-5.
- Mehedi, M., Falzarano, D., Seebach, J., Hu, X., Carpenter, M. S., Schnittler, H.-J., et al. (2011). A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J. Virol.* 85, 5406–14. doi:10.1128/JVI.02190-10.
- Mühlberger, E., Trommer, S., Funke, C., Volchkov, V., Klenk, H.-D., and Becker, S. (1996). Termini of all mRNA species of Marburg virus: Sequence and secondary structure. *Virology* 223, 376–380. Available at: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed4&NEWS=N&AN=1996293636>.
- Mühlberger, E., Weik, M., Volchkov, V. E., Klenk, H. D., and Becker, S. (1999). Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J. Virol.* 73, 2333–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9971816> [Accessed July 17, 2019].
- Neumann, G., Watanabe, S., and Kawaoka, Y. (2009). Characterization of Ebolavirus regulatory genomic regions. *Virus Res.* 144, 1–7. doi:10.1016/j.virusres.2009.02.005.
- Pigott, D. C. (2005). Hemorrhagic Fever Viruses. *Crit. Care Clin.* 21, 765–783. doi:10.1016/j.ccc.2005.06.007.
- Rougeron, V., Feldmann, H., Grard, G., Becker, S., and Leroy, E. M. (2015). Ebola and Marburg haemorrhagic fever. *J. Clin. Virol.* 64, 111–119. doi:10.1016/j.jcv.2015.01.014.
- Sanchez, A., Kiley, M. P., Holloway, B. P., and Auperin, D. D. (1993). Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res.* 29, 215–240. doi:10.1016/0168-

1702(93)90063-S.

- Schlereth, J., Grünweller, A., Biedenkopf, N., Becker, S., and Hartmann, R. K. (2016). RNA binding specificity of Ebola virus transcription factor VP30. *RNA Biol.* 13, 783–798. doi:10.1080/15476286.2016.1194160.
- Shabman, R. S., Jabado, O. J., Mire, C. E., Stockwell, T. B., Edwards, M., Mahajan, M., et al. (2014). Deep Sequencing Identifies Noncanonical Editing of Ebola and Marburg Virus RNAs in Infected Cells. 5, 1–11. doi:10.1128/mBio.02011-14.Editor.
- Shu, T., Gan, T., Bai, P., Wang, X., Qian, Q., Zhou, H., et al. (2019). Ebola virus VP35 has novel NTPase and helicase-like activities. *Nucleic Acids Res.* 47, 5837–5851. doi:10.1093/nar/gkz340.
- Sugita, Y., Matsunami, H., Kawaoka, Y., Noda, T., and Wolf, M. (2018). Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex at 3.6 Å resolution. *Nature*. doi:10.1038/s41586-018-0630-0.
- Tomescu, A. I., Robb, N. C., Hengrung, N., Fodor, E., and Kapanidis, A. N. (2014). Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc. Natl. Acad. Sci.* 111, E3335–E3342. doi:10.1073/pnas.1406056111.
- Volchkov, V. E., Volchkova, V. A., Chepurinov, A. A., Blinov, V. M., Dolnik, O., Netesov, S. V., et al. (1999). Characterization of the L gene and 5' trailer region of Ebola virus. *J. Gen. Virol.* 80, 355–362. doi:10.1099/0022-1317-80-2-355.
- Wan, W., Kolesnikova, L., Clarke, M., Koehler, A., Noda, T., Becker, S., et al. (2017). Structure and assembly of the Ebola virus nucleocapsid. *Nat. Publ. Gr.* 551, 394–397. doi:10.1038/nature24490.
- Weik, M., Enterlein, S., Schlenz, K., and Muhlberger, E. (2005). The Ebola Virus Genomic Replication Promoter Is Bipartite and Follows the Rule of Six. *J. Virol.* 79, 10660–10671. doi:10.1128/jvi.79.16.10660-10671.2005.
- Weik, M., Modrof, J., Klenk, H.-D., Becker, S., and Muhlberger, E. (2002). Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J. Virol.* 76, 8532–9. doi:10.1128/JVI.76.17.8532.
- Yang, D., and Leibowitz, J. L. (2015). The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res.* doi:10.1016/j.virusres.2015.02.025.



## FIGURE LEGENDS

**Fig. 1:** Secondary structure formation potential at (A) transcription start regions of EBOV genes at the genomic (top) and antigenomic (bottom) RNA level and (B) sequence and features of the EBOV 3'-leader promoter. **(A)** Genomic sequence elements required for transcription (re)initiation are shown in cyan at the top; transcription is initiated opposite to the 3'-terminal C residue in the transcription start signal (TSS). Schematic white boxes in the center mark the open reading frames for proteins NP, VP35, VP40, GP, VP30, VP24 and L; light gray boxes indicate 5'- and 3'-UTRs, with dark gray areas depicting the position of the predicted secondary structures on the genomic (top) or antigenomic/mRNA level (bottom); p, leader and trailer promoters (Calain et al., 1999); mRNAs of the 7 EBOV genes are shown as horizontal colored lines with terminal dots indicating their 5'-caps. Secondary structures are the minimum free energy (MFE) structures (predicted by RNAfold using the default parameters; <http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>).  $\Delta G$  values are depicted in red above (vRNA level) or below (mRNA level) the structures. **(B)** Validated secondary structures forming in naked leader RNA (Schlereth et al., 2016; Weik et al., 2002, 2005). Proposed promoter elements 1 and 2 (PE1, PE2) are shown in green letters and the 3'-U residues of UN<sub>5</sub> hexamers in PE1 and PE2 are highlighted in pink in the genomic RNA; orange nucleotides mark the spacer region between the transcription start sequence (TSS, in cyan) and PE2. The distance from nt -51 to -80 ( $5 \times 6 = 30$  nt) was recently defined as measure for hexamer phasing between PE1 and PE2 (Bach et al., 2019). Nucleotide numbering of the genomic RNA starts at the 3'-terminal nt (position -1) that is complementary to position 1 (5'-terminus) of the antigenome. The 3'-terminal G of the genome is shown as small letter to consider the recent finding that this nucleotide is not essential for initiation, as the EBOV RNA polymerase initiates RNA synthesis at the C residue preceding the 3'-terminal G residue (Deflubé et al., 2019).

**Fig. 2:** **(A)** Illustration of EBOV 3'-leader minigenome constructs (replication-competent) in which the NP (wt) HP (turquoise-shaded area) consisting of the TSS and the orange spacer sequence was replaced with HP structures derived from the transcription reinitiation sites of the VP40, L and VP24 genes; in the case of the latter two, 1 nt (L) or 2 nt (VP24) were additionally inserted at the indicated locations to restore hexamer phasing between nt -51 and -80 (Bach et al., 2019). For further details, see Fig. 1B. **(B)** Relative reporter activity of replication-competent minigenome (RC MG) constructs specified in panel A. The extension of the spacer region relative to the parental NP (wt) variant is given above the bars. Activity values ( $\pm$  standard error of the mean, SEM) were normalized to the NP (wt) construct and are based on at least 3 biological replicates with

2 or 3 technical replicates each. - L, background control (transfection without the plasmid encoding the polymerase L); white columns: native 5'-UTR sequences; blue columns: 5'-UTR sequences engineered to obey the “rule of 6”.

**Fig. 3:** (A) Predicted MFE structures (RNAfold) on the vRNA (top) and mRNA (bottom) level of engineered L and VP24 constructs ( $\Delta G$  values in red). The individual spacer extensions (in nt) are given at the top; del., nt deletions; ins., nt insertions. (B) Corresponding MFE structures of the VP40 constructs with increasing spacer extensions between PE1 and PE2. (C) Reporter gene activities measured for the replication-competent minigenome (RC MG) constructs specified in panels A and B. Values were normalized to the NP (wt) 3'-leader as control (100 %). - L, background control (specified in the legend to Fig. 2B). Mean values ( $\pm$  SEM) are based on three independent experiments with three technical replicates each. White columns: reference constructs; blue columns: engineered VP24 HP constructs; pink columns: engineered VP40 HP constructs.

**Fig. 4:** Analysis of a 3'-leader construct with a 12-nt deletion in the spacer region ( $\Delta 5'$ -spacer). (A) Predicted MFE structure of the NP (wt) TSR and the corresponding region of the  $\Delta 5'$ -spacer variant carrying a deletion of 12 spacer nucleotides. The latter variant is predicted to be devoid of any secondary structure on the vRNA and mRNA level. For the nucleotide color code on the vRNA level, see legend to Fig. 1B. (B) Luciferase reporter activities measured for the  $\Delta 5'$ -spacer construct, the NP (wt) construct and the NP (wt) construct in the absence of polymerase L (background control, see legend to Fig. 2B). RC MG, replication-competent minigenome. Mean values ( $\pm$  SEM), normalized to the native NP (wt) leader construct, are based on 3 biological replicates with 3 or 4 technical replicates each; \*\*\*\* $p < 0.00001$ ; (unpaired  $t$  test, two-tailed). (C-E) Corresponding levels of viral mRNA, cRNA and vRNA measured by a 2-step strand-specific qRT-PCR of replication-competent minigenome samples using the same cells as in panel B; normalized to the NP (wt) construct. For details, see Bach et al. (2019). Mean  $2^{-\Delta\Delta CT}$  values ( $\pm$  SEM) were derived from at least three independent experiments with 3 or 4 technical replicates each; \* $p < 0.05$ ; \*\*\* $p < 0.0001$ ; \*\*\*\* $p < 0.00001$ ; (unpaired  $t$  test, two-tailed; a corresponding Mann Whitney test was used in case of data not normally distributed).

**Fig. 5:** Predicted MFE structures (RNAfold) of minigenome constructs encoding the NP (wt) HP, its stabilized derivative NP G<sub>-72</sub>, the native VP40 HP and its incrementally stabilized derivatives VP40<sub>6 to 1</sub>; predicted  $\Delta G$  values (in red) are given for the vRNA strand (top) and the encoded mRNA (bottom); del., single nt deletions; pink residues,

point mutations; green residues, single nt insertions; the pink line at the genomic VP40 HP marks the GE signal of the preceding VP35 gene.

**Fig. 6:** Impact of NP HP stabilization and progressive rigidification of the VP40 HP on minigenomic viral transcription and replication. **(A)** Luciferase activities of replication-competent minigenomes (RC MG) carrying the NP and VP40 variants illustrated in Fig. 5. Activity values are given in % relative to the native 3'-leader [NP (wt) = 100 %]. - L, background control (see legend to Fig. 2B). Mean values ( $\pm$  SEM) are based on 3 biological replicates with at least 3 technical replicates each; \* $p < 0.05$ ; \*\*\* $p < 0.0001$ ; \*\*\*\* $p < 0.00001$  (unpaired  $t$  test, two-tailed). **(B-D)** Corresponding levels of viral mRNA, cRNA and vRNA measured by a 2-step strand-specific qRT-PCR of replication-competent minigenome samples using the same cells as in panel A. For details, see Bach et al. (2019). Mean values ( $\pm$  SEM) were derived from 3 independent experiments with at least 3 technical replicates each. \* $p < 0.05$ ; \*\* $p < 0.001$ ; \*\*\*\* $p < 0.00001$ ; (unpaired  $t$  test, two-tailed, a corresponding Mann Whitney test was used in case of data not normally distributed).

FIGURES

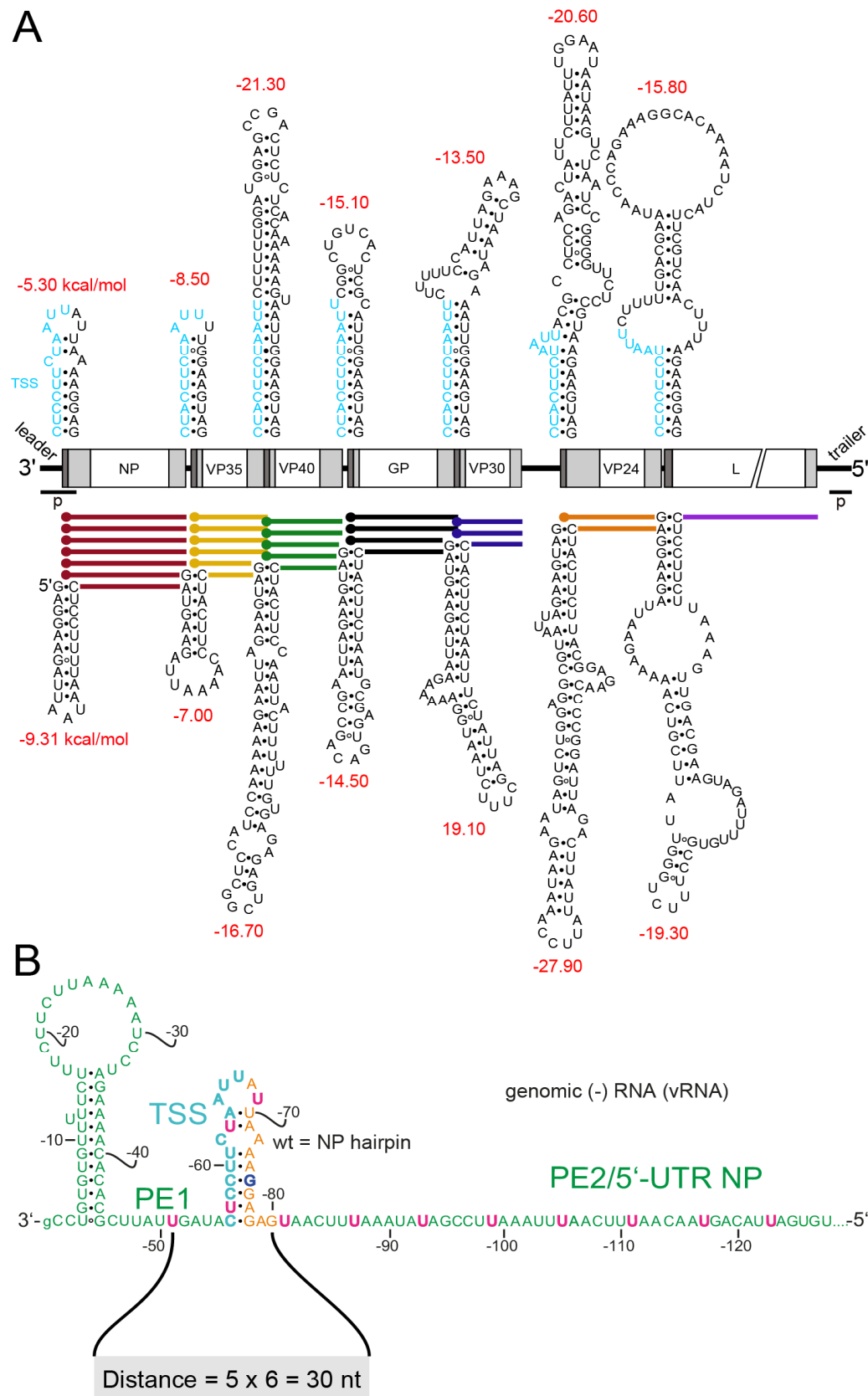


Fig. 1

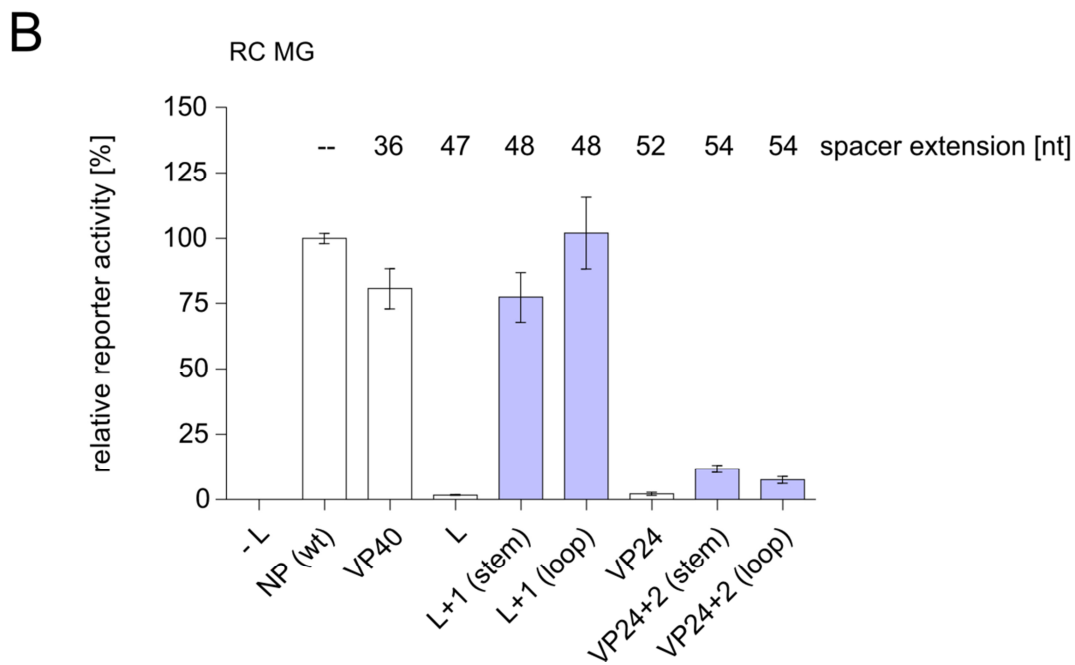
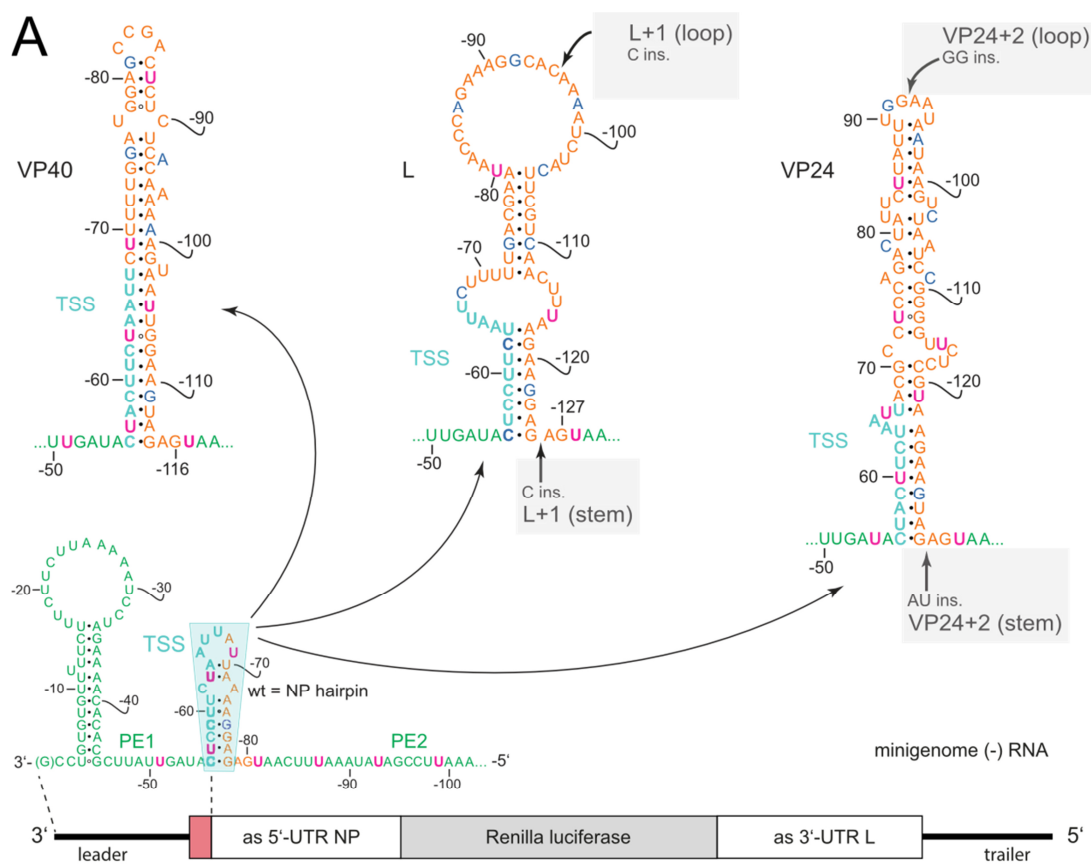


Fig. 2

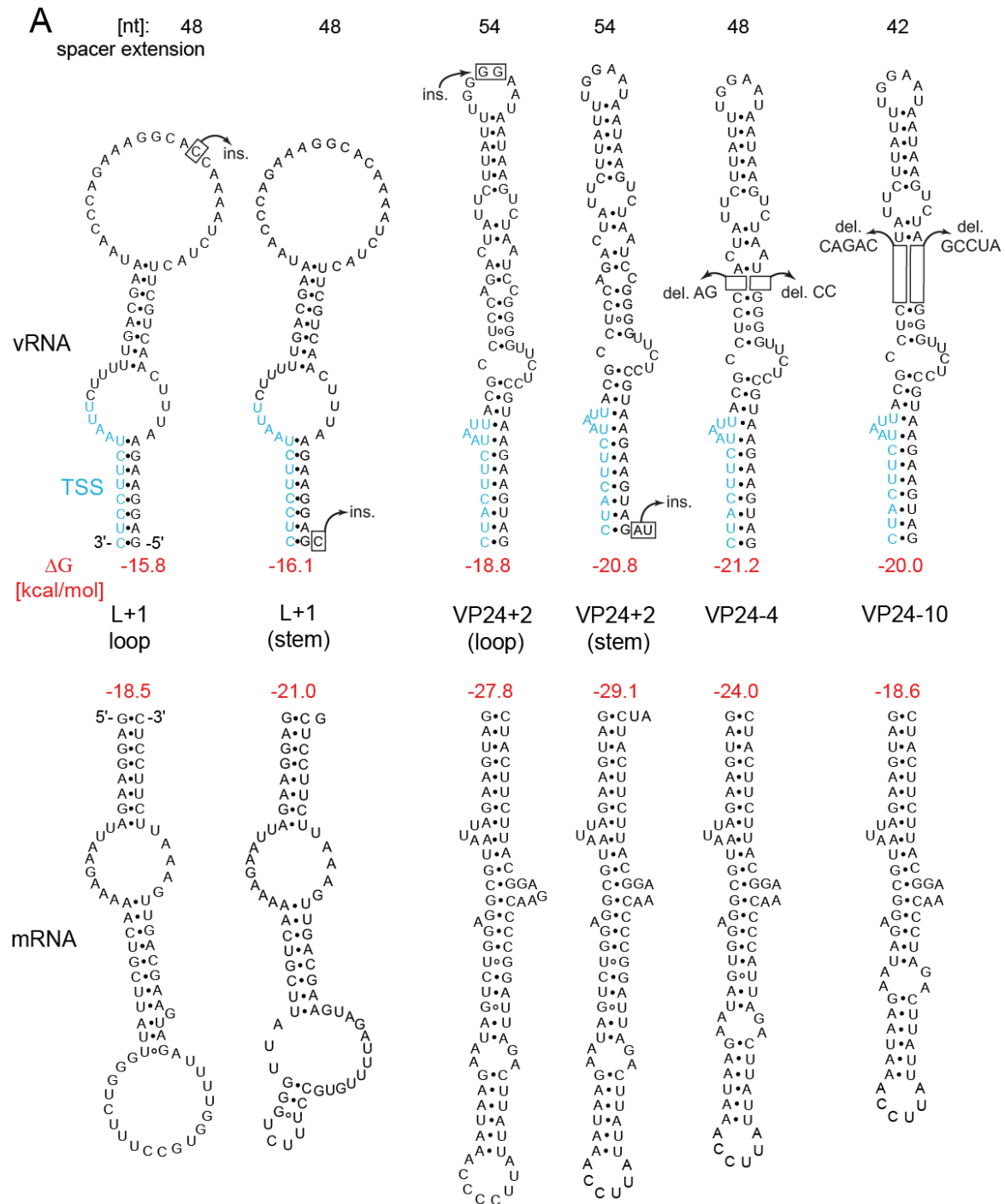


Fig. 3A

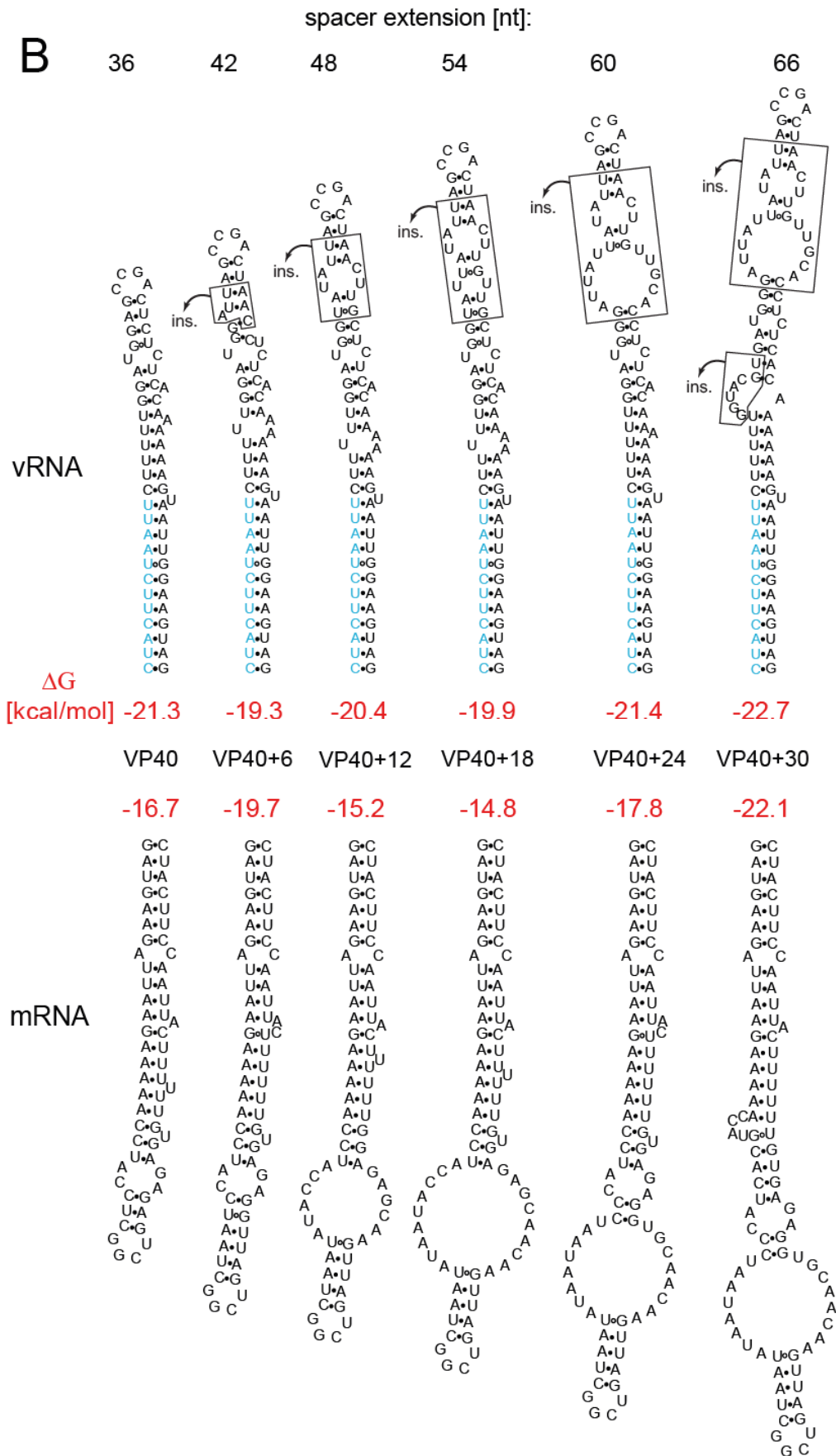


Fig. 3B

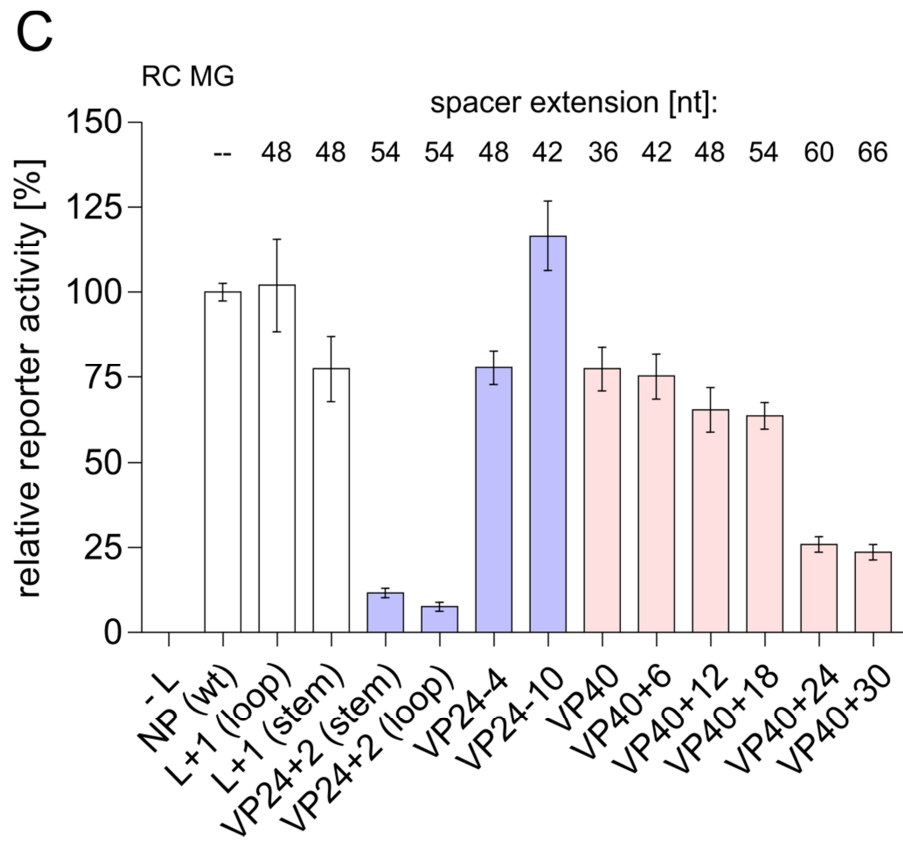


Fig. 3C



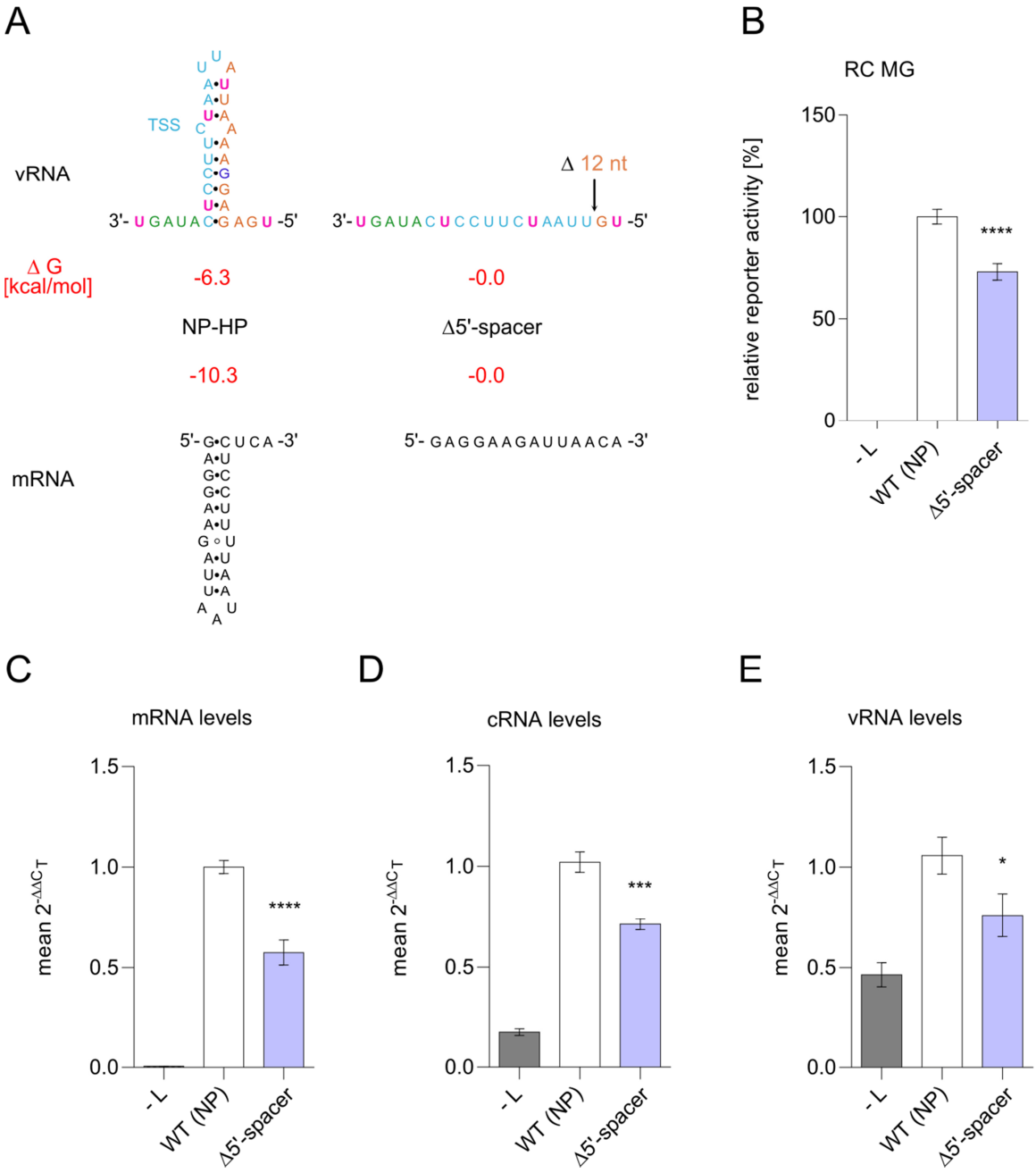


Fig. 4

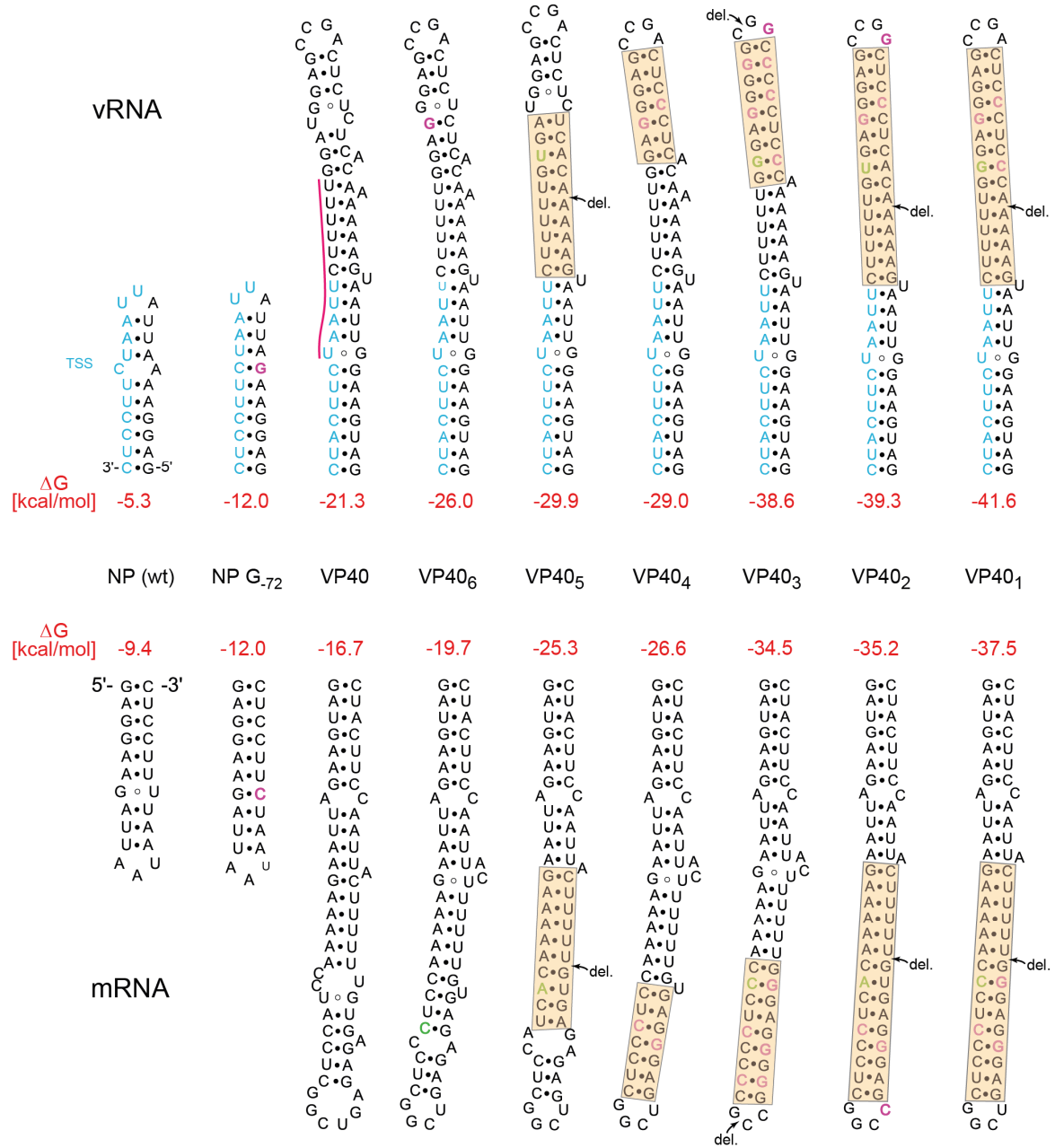
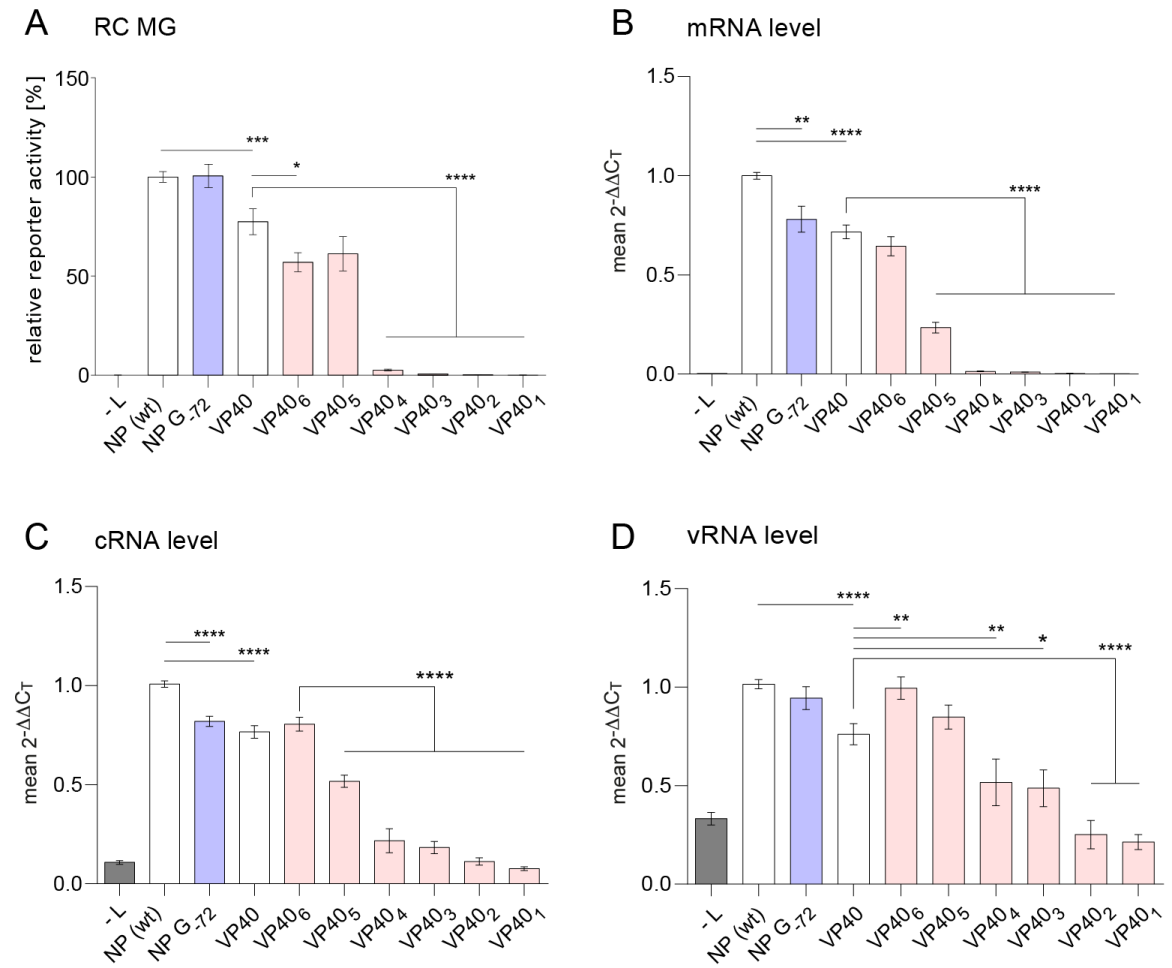


Fig. 5



**Fig. 6**

## **SUPPLEMENTARY MATERIAL**

### **Secondary structure formation influences EBOV transcription initiation and replication in a length- and stability-dependent manner**

Simone Bach<sup>1</sup>, Nadine Biedenkopf<sup>2</sup>, Stephan Becker<sup>2, #</sup>, Roland K. Hartmann<sup>1, #</sup>

<sup>1</sup>Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35037 Marburg, Germany

<sup>2</sup>Institut für Virologie, Philipps-Universität Marburg, Hans-Meerwein-Str. 2, 35043 Marburg

#corresponding author:

E-mail:       roland.hartmann@staff.uni-marburg.de

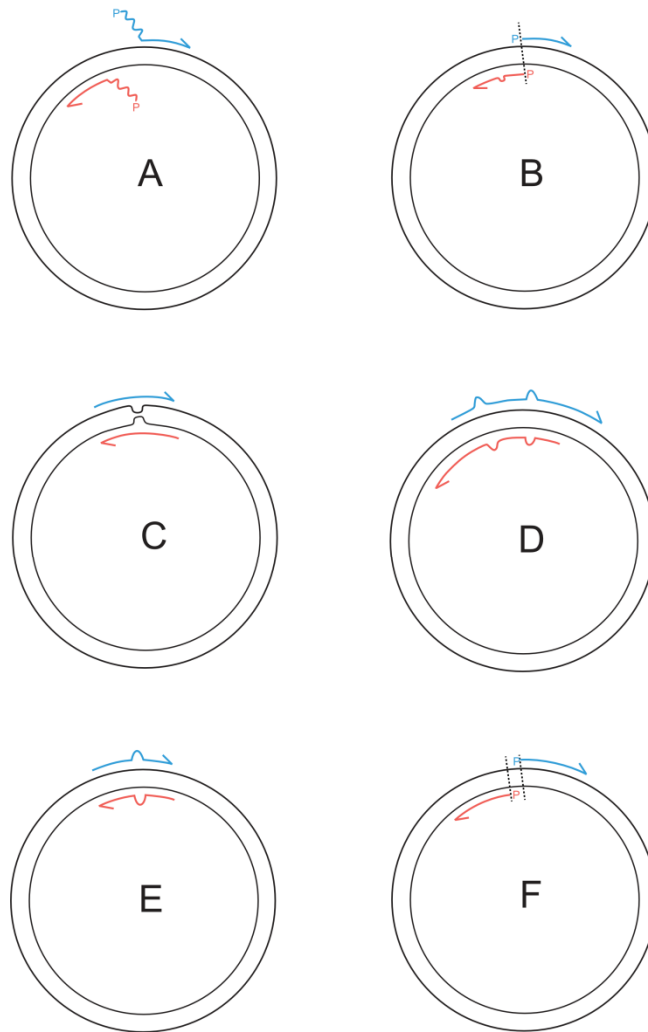
**Table S1:** Primers used in this study for the construction of mutant minigenomes. The different PCR-based strategies for site-directed mutagenesis are illustrated in Fig. S1 (except for the classical cloning using restriction enzymes). The EBOV wt minigenome pANDY 3E5E is termed p3E5E\_Rluc\_RC; RC indicates the replication-competent minigenomes. Dashes in primer sequences mark nucleotide deletions, boldface underlined nucleotides substitutions or insertions.

**Tab. 37: Cloning strategies and primers manuscript 2.**

no.	name p3E5E_	Cloning strategy: site-directed mutagenesis / plasmid template	Primer fwd name	primer fwd (5' to 3')	primer rev name	primer rev (5' to 3')
Overhang/inside-out primer insertion/substitution mutagenesis						
1	VP40+12	2-step mutagenesis of p3E5E_VP40 (Bach et al., 2019)	pS_0120	<u>CAATCAGCCGAGGT</u> AGGTTTTCTTAATC	pS_0121	<u>AACGAGAGT</u> GTTTTTT CATTAACTTCATCTC ATTG
			pS_0122	<u>TATGGTAGGTTTTTC</u> TTAATCTTCATCATA GTTATTGCGACAC	pS_0123	<u>TAATCGGCTGATTGA</u> ACGAGAGTGTGTTTTTC ATTAACC
2	VP40 <sub>3</sub>	mutagenesis of p3E5E_VP40	pS_168	<u>C—GGGGGAGGGT—T</u> TTTCTTAATCTTCAT CATAGTTATTGCGAC	pS_167	<u>CCGGGGGAGGGTTTT</u> TTCATTAACTTCATC TCATTGAAATTTATAT CG
3	VP40 <sub>2</sub>	mutagenesis of p3E5E_VP40	pS_0170	<u>CCGAGGGAGT</u> GTTTT TTCTTAATCTTCATC ATAGTTATTGCGACA C	pS_0169	CCGAGGGAGTGTGTTTT TCATTAACTTCATCT CATTGAAATTTATATC GG
4	VP40 <sub>1</sub>	mutagenesis of p3E5E_VP40	pS_172	<u>CCGAGGGAGGGTTTT</u> TTCTTAATCTTCATC ATAGTTATTGCGACA C	pS_171	<u>CTGAGGGAGGGT—TT</u> TTCATTAACTTCATC TCATTGAAATTTATAT CGG
Substitution mutagenesis						
14	VP40 <sub>5</sub>	mutagenesis of p3E5E_VP40	pS_241	CACTCTCTCAGCCG AGGTAGTGTGTTTTCT TAATCTTCATC	pS_242	TTTTTCATTAACTTC ATCTCATTGAAATTTA TATCGGAATTTAAATT G
Complementary primer mutagenesis						
5	VP24-4	2-step deletion mutagenesis of p3E5E_VP24 (Bach et al., 2019)	pS_179	GAATGCCTCTTGGG G—TAATCTGAATAA TAAGG	pS_180	CCTTATTATTCAGATT A—CCCCAAGAGGCA TTC
			pS_181	GGTTTATTCTTATCA —CCTCCGCATTAAT CTTCATC	pS_182	GATGAAGATTAATGC GGAGG—TGATAAGA ATAAACC
6	VP40+6	mutagenesis of p3E5E_VP40	pS_185	CACTCTCCAATCAG CCGATTAGGTAGGT TTTTCTTAATC	pS_186	CCTACCTAATCGGCT GATTGGAGAGTGTGTT TTTCATTAAAC
7	VP40+18	mutagenesis of p3E5E_VP40+12 (Bach et al., 2019)	pS_183	<u>CACTCTCGTTGTTCA</u> ATCAGCCGATTATAT <u>TATGGTAGGTTTTTC</u>	pS_184	GAAAAACCTACCATAA TATAATCGGCTGATTG AAC <u>AAC</u> GAGAGTG

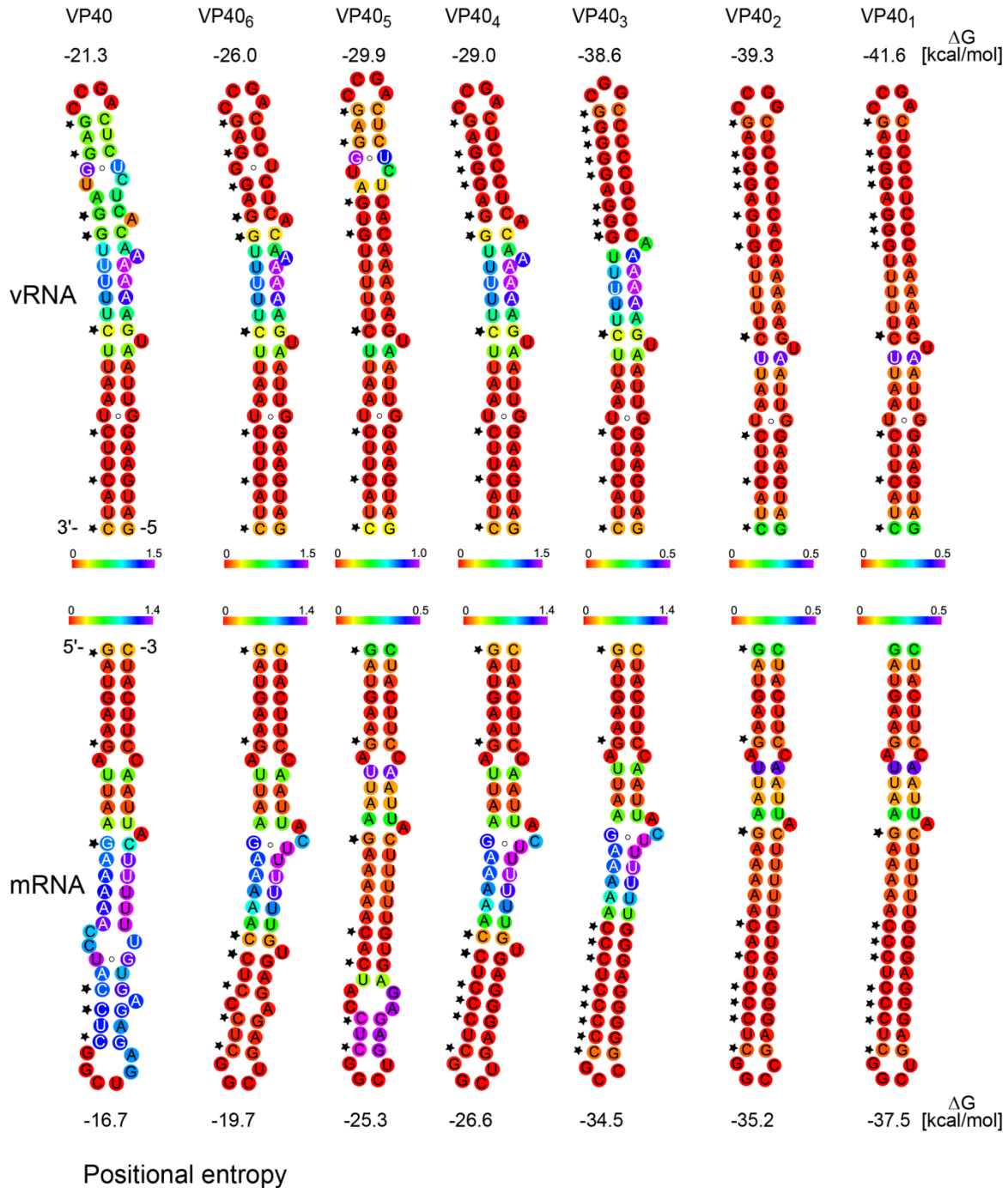
no.	name p3E5E_	Cloning strategy: site-directed mutagenesis / plasmid template	Primer fwd name	primer fwd (5' to 3')	primer rev name	primer rev (5' to 3')
Complementary primer mutagenesis						
8	VP40+24	mutagenesis of p3E5E_VP40+18	pS_196	GAAGAAACACTCTC CACGTTGTTCAATCA GCCGATTATATTATT AGGGTAGG	pS_197	CCTACCCTAATAATAT AATCGGCTGATTGAA CAACGTGGAGAGTGT TTTTTC
9	VP40+30	mutagenesis of p3E5E_VP40+24	pS_198	GCCGATTATATTATT AGGGTAGTGCATGG TTTTCTTAATC	pS_199	GATTAAGAAAAACCAT GCACTACCGTAATAAT ATAATCGGC
10	VP40 <sub>4</sub>	mutagenesis of p3E5E_VP40	pS_223	GTTAATGAAAAACA CTCCCTCAGCCGAG GGAGGTTTTCTTAA TCTTC	pS_224	GAAGATTAAGAAAA CCTCCCTCGGCTGAG GGAGTGTTTTTCATT AAC
11	VP40 <sub>6</sub>	mutagenesis of p3E5E_VP40	pS_256	CTCTCTCAGCCGAG GGAGGTTTTCTTAA TCTTC	pS_257	GAAGATTAAGAAAA CCTCCCTCGGCTGAG AGAG
Inside-out primer deletion mutagenesis						
12	Δ5'-spacer	mutagenesis of p3E5E_Rluc_RC	pS_0100	TTAATCTTCCTCATA GTTATTCGCACAC	pS_0101	CATTGAAATTTATATC GGAATTTAAATTGAAA TTG
13	VP24-10	2-step mutagenesis of p3E5E_VP24	pS_0116	ATCTGAATAATAAGG TTTATTCTTATCAGA CCTCCG	pS_0117	CCCAAGAGGCATTCT TCATCTCATTG
			pS_0118	CTCCGCATTAATCTT CATCATAGTTATTCT C	pS_0119	ATAAGAATAAACCTTA TTATTCAGATCCCAAG AGGC

**Fig. S1:** PCR-based strategies (essentially performed as previously described, Li et al., 2009) that were used for the construction of mutant minigenomes. **(A)** Overhang/inside-out primer insertion mutagenesis. **(B)** Substitution mutagenesis. **(C, D, E)** Complementary primer mutagenesis; deletions were introduced according to scheme C and insertions or substitutions according to scheme D and E. **(F)** Inside-out primer deletion mutagenesis. For more information, see also the NEB web site for the Q5 Site-Directed Mutagenesis Kit (<https://www.neb-online.de/en/pcr-and-dna-amplification/q5-site-directed-mutagenesis-kit/>). In approaches A, B and F, the entire plasmid is amplified with 5'-phosphorylated primers that introduce the desired substitutions/insertions/deletions, followed by circularization of PCR products and template removal by *DpnI* treatment before bacterial transformation. In approaches C, D and E, phosphorylation of 5'-ends and ligation are carried out by bacterial enzymes after DNA transformation.



**Fig. S1**

**Fig. S2:** Positional entropy of the VP40 HP variants shown in Fig. 5 of the main manuscript to highlight stretches of rigid duplex structure. Asterisks mark G:C base pairs; circle colors indicate low (red), high (violet) and intermediate (increasing from orange to blue) positional entropy; range bars of positional entropy are indicated in the center below or above each structure, respectively.



**Fig. S2**



**Reference**

- Li, D., Willkomm, D. K., and Hartmann, R. K. (2009). Minor changes largely restore catalytic activity of archaeal RNase P RNA from *Methanothermobacter thermoautotrophicus*. *Nucleic Acids Res.* 37, 231–242. doi:10.1093/nar/gkn915.

### 4.3 Detektion und Charakterisierung von EBOV-*leader*RNAs

Die Transkription der nicht-segmentierten, Negativstrang (NNS)-RNA-Viren folgt einem Stopp-Start-Mechanismus, der aller Voraussicht nach ausschließlich am äußersten 3'-Ende des Virusgenoms initiiert werden kann. Es wird angenommen, dass die virale Polymerase vor der Initiation der Transkription am ersten NP-Gen zunächst ein kurzes, 5'-Kappen-looses, nicht polyadenyliertes und nicht enkapsidiertes Antisense-Transkript synthetisiert, das komplementär zum 3'-Ende des Genoms ist. In der vorliegenden Studie wurden RNA-Sequenzierungen, Northern Blot- und qRT-PCR-Analysen von Ebola-Virus (EBOV)-infizierten oder EBOV-Minigenom-transfizierten Zellen durchgeführt. Diese bestätigen die Synthese abortiver EBOV-*leader*RNAs mit einer heterogenen Länge von ca. 60 – 80 Nt. Wir können zeigen, dass die virale *leader*RNA-Synthese der „Rule of 6“ unterliegt und spezifisch am zweiten Genomnukleotid (Nt -2) initiiert wird. Jedoch scheint das erste Genomnukleotid (G-1) für die Steigerung der Transkriptionsinitiation (~ 4-fach) am NP-Transkriptionsstartsignal (NP-TSS) wichtig zu sein. 24 h nach EBOV-Infektion, bzw. 48 h nach Minigenom-Transfektion, liegen (NP) mRNA:leaderRNA-Konzentrationsverhältnisse von etwa 9:1 (EBOV-Infektion) und 68:1 (Minigenom-Transfektion) vor. In Minigenomassays bleiben sowohl die exklusive *leader*RNA-Initiationsstelle an Nt -2 als auch die *leader*RNA-Terminationsregion weitgehend unbeeinflusst von RNA-Sequenz- und -Struktur-variationsen, einer verkürzten Spacer-Länge zwischen PE1 und PE2 oder der Abwesenheit von VP30. Darüber hinaus sind die *leader*RNA-Spiegel in Anwesenheit von VP30 reduziert. Die Ergebnisse sind mit der Hypothese vereinbar, dass der L/VP35-Polymerasekomplex Leader-Transkripte ohne die Einbindung des Transkriptionsaktivators VP30 synthetisiert. Dieser assoziiert erst während der Transkriptionsinitiation am ersten NP-TSS mit dem Polymerasekomplex. Es ist folglich denkbar, dass *leader*RNAs abortive Antigenom-RNAs und keine obligaten Vorlauf-Produkte der Transkriptionsinitiation am ersten NP-Gen sind. Darüber hinaus diskutieren wir auf der Grundlage bioinformatischer Strukturvorhersagen sowie erster experimenteller Ergebnisse die mögliche Funktion von EBOV-Leader-Transkripten als Effektoren der RIG-I-vermittelten angeborenen Immunantwort.

## **Ebola virus transcriptome analysis reveals insights into transcriptional regulation**

Simone Bach<sup>1</sup>, Julia Schlereth<sup>1</sup>, Nadine Biedenkopf<sup>2</sup>, Stephan Becker<sup>2</sup>, Marcus Lechner<sup>3</sup>, Roland K. Hartmann<sup>1,#</sup>

<sup>1</sup> Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35037 Marburg, Germany

<sup>2</sup> Institut für Virologie, Philipps-Universität Marburg, Hans-Meerwein-Str. 2, 35043 Marburg

<sup>3</sup> Center for Synthetic Microbiology, Hans-Meerwein-Str. 6, 35032 Marburg

#corresponding author:

E-mail:       roland.hartmann@staff.uni-marburg.de

---

**ABSTRACT**

Transcription of non-segmented negative strand (NNS) RNA viruses follows a stop-start mechanism and is thought to be initiated exclusively at the genome's very 3' end, producing a short uncapped, non-polyadenylated and non-encapsidated transcript antisense to the 3'-leader before initiating transcription at the first NP gene. Here, we performed RNA-Seq of Ebola virus (EBOV)-infected or EBOV minigenome-transfected cells as well as Northern Blot and qRT-PCR analyses, demonstrating the synthesis of abortive *leader*RNAs with heterogeneous lengths (~ 60-80 nt) in the EBOV system. We show that viral *leader*RNA synthesis follows the "rule of 6" and is specifically initiated opposite genome nt -2. However, the first genome nt (G<sub>-1</sub>) needs to be present for efficient transcription initiation at the transcription start site. We determined a steady-state NP mRNA:*leader*RNA ratio of ~ 9:1 at 24 h after EBOV infection, and this ratio was higher (68:1) for minigenome-transfected cells (48 h post transfection). *Leader*RNA initiation at nt -2 and the range of termination sites remained basically unchanged upon RNA sequence and structure variation, shortage of the spacer between promoter elements 1 and 2, or in the absence of VP30 in minigenome assays. Our results further indicate that the presence of VP30 suppresses *leader*RNA synthesis. Our findings are consistent with EBOV *leader*RNAs being abortive antigenomic RNA products rather than obligatory pre-products of transcription initiation at the first NP gene. Finally, we provide evidence arguing against a potential role of EBOV leader transcripts as effector molecules in innate immunity.

## INTRODUCTION

The Ebola virus (EBOV), a member of the *Filoviridae* family in the order *Mononegavirales*, is the causative agent of the Ebola virus disease with high fatality rates of up to 90 % in humans in past outbreaks (Garske et al., 2017). Its non-segmented negative sense (NNS) genome (Fig. 1A), 19 kb in length, serves as template for nine mRNAs that are translated into seven structural proteins and two nonstructural proteins. The latter two are generated by co-transcriptional editing of the glycoprotein (GP) mRNA that provides two additional versions of the GP protein, soluble (sGP) and small soluble GP (ssGP; Mehedi et al., 2011; Rougeron et al., 2015). The seven open reading frames are flanked by 5' and 3' untranslated regions (UTRs). The terminal regions of the EBOV genome encode promoters for initiation of RNA synthesis by the viral polymerase complex as well as signals important for encapsidation.

The genomic 3'-leader harbors the EBOV replication and transcription promoter (approx. nt -1 to -156) and the antigenomic replication promoter is encoded in the 5'-terminal 176 nt of the RNA genome (Calain et al., 1999; Sanchez et al., 1993; Volchkov et al., 1999). The 3'-leader harbors a transcription initiation sequence, termed gene start (GS) signal (nt -56 to -67), and the transcription promoter is assumed to be located immediately 3' of this GS signal (Mühlberger et al., 1998) but has not been pinpointed to specific nucleotide positions yet. However, there is evidence that genome nucleotides -54 and -55 are obligatory to efficient transcription initiation at this first GS signal directing synthesis of the first nucleoprotein (NP) mRNA (Bach et al., 2019a). The filoviral replication promoter was shown to be bipartite, a feature that is only shared with the *paramyxoviridae* among NNS viruses (Calain and Roux, 1993; Kolakofsky, 2016; Le Mercier and Kolakofsky, 2019; Tapparel et al., 1998). In the 3'-leader, the first promoter element (PE1) is located in the 3'-terminal 55 nt of the genome, preceding the GS and a spacer region which separates PE1 from the second promoter element (PE2) spanning positions -81 and -128 (Fig. 1B). PE2 harbors eight consecutive 3'-UN<sub>5</sub> hexamers. Mutational analysis suggested that base identities at positions -10 and -13 as well as -44 to -55 are crucial for efficient replication (Weik et al., 2005). Deletion of the terminal 55 or 56 nt of the trailer promoter prevented rescue of full-length EBOV, suggesting that this region is important for replication and/or encapsidation (Hoenen et al., 2010; Sztuba-Solinska et al., 2016). Likewise, removal of the terminal 25 nt of the trailer promoter reduced replication to a single round, thus abrogating *de novo* vRNA synthesis (Deflubé et al., 2019; Olsen et al., 2018). Furthermore, hexamer phasing ("rule of 6") between PE1 and PE2 of the 3'-leader promoter has been shown to play a key role in both, productive replication and transcription (Weik et al., 2005; Bach et al., 2019a).

Transcription and replication occurs at the helical viral nucleocapsid consisting of NP filaments that enwrap the template RNA (Wan et al., 2017). In the nucleocapsid, 6 nt of the RNA are bound per NP molecule (Sugita et al., 2018; Wan et al., 2017). This makes it likely that hexamer phasing in the 3'-leader promoter is functionally linked to NP coverage.

While replication of the NP-RNA template is conducted by the viral polymerase L and its cofactor VP35, EBOV transcription additionally requires protein VP30 (Mühlberger et al., 1999). In the presence of VP30, viral replication is concurrently down-regulated (Biedenkopf et al., 2013). VP30's capability to enhance transcription is dependent on RNA binding, interaction with VP35 (mediated by RNA) and its phosphorylation status. While non- or hypo-phosphorylated VP30 favors transcription, phosphorylation of VP30 enhances its binding affinity to NP which impairs VP35 binding and presumably leads to a release of VP30 from the viral transcription complex to favor viral replication (Biedenkopf et al., 2016b, 2016a; Kruse et al., 2018; Schlereth et al., 2016). Filoviral mRNA synthesis is assumed to follow a stop-start mechanism regulated by highly conserved GS and gene end (GE) signals that are recognized by the viral RNA-dependent RNA polymerase complex (RdRp). Termination of the upstream mRNA transcript was shown to be a prerequisite for efficient initiation at the downstream GS (Brauburger et al., 2016). Similar to other viruses of the *Mononegavirales*, the RdRp is thought - after termination of the upstream gene - to either reinitiate transcription at the downstream GS or to dissociate from the genome template and reenter the genome exclusively at the 3' end for new transcription initiation at the NP gene (revised in Hume and Mühlberger, 2019; Mühlberger, 2007). This mechanism explains the trend that viral mRNA levels decline from the first (NP) to the last (L) gene, which has gained support from recent transcriptome analyses of cells infected with various filoviruses including EBOV (Albariño et al., 2018; Shabman et al., 2014), as schematically illustrated in Fig. 1A. Interestingly, EBOV antigenome synthesis was recently shown to be initiated not at the genome very 3' end but at genome pos. -2. Thus, the genome's 3'-terminal nucleotide is thought to be added to newly synthesized vRNA genomes in a non-templated manner (Deflubé et al., 2019).

According to present knowledge, the filoviral polymerase can only initiate RNA synthesis from the 3'-end of the RNA template (single-entry model). It has been shown for other NNS viruses also utilizing a sequential stop-start-transcription mechanism that the polymerase initiates RNA synthesis at the genome 3'-end, either uniformly at the 3'-terminal nucleotide (Sendai virus, SeV) or at the 3'-terminus (pos -1) as well as at pos. -3 (respiratory syncytial virus, RSV). In RSV, polymerases initiating RNA synthesis

2 nt upstream of the 3'-end are thought to be unable to adopt a stable elongation mode and thus release an abortive *leader*RNA of ~ 25 nt; they then scan the template RNA for the next GS signal to initiate mRNA transcription (Fearn and Plemper, 2017; Noton and Fearn, 2015; Ogino and Green, 2019; Whelan et al., 2004). The model further considers that replicative RNAs initiated at the 3'-terminal nucleotide may be aborted as well when the amounts of nucleoprotein are insufficient for encapsidation (Tremaglio et al., 2013). This raises the question as to whether EBOV shares at least some of these mechanistic features with RSV and other NNS viruses. For filoviruses, however, the existence of leader transcripts has not been proven yet. In the present study, we were able to detect such abortive *leader*RNAs in EBOV-infected cells and EBOV minigenome systems, providing new mechanistic insights into polymerization initiation in filoviruses. Abortive *leader*RNAs in total RNA of EBOV-infected HuH7 cells or in total RNA of minigenome (MG)-transfected HEK293 cells were detected and quantified by RNA-Seq, Northern Blot as well as qRT-PCR analyses. We further investigated whether *leader*RNAs may act as effectors in innate immunity.

## RESULTS

### Transcriptome analysis of EBOV-infected HuH7 cells reveals synthesis of abortive *leader*RNAs

We performed a transcriptome (RNA-Seq) analysis of EBOV-infected HuH7 cells either enriching for poly(A) RNA to detect NP mRNAs or enriching for small RNAs < 200 nt to detect potential *leader*RNAs. Indeed, considerable amounts of short transcripts antisense to the genome 3'-leader sequence were identified in the small RNA library. The majority (~ 92 %) was 60-80 nt in length (Fig. 2A), thus resulting from termination shortly after the transcription start sequence (TSS) and before the PE2 region (Fig. 1B). Hence, the RNA-Seq data provide evidence for the production of abortive *leader*RNAs by the EBOV polymerase complex. The termination sites coincide with the region of the NP hairpin (HP) structure. The NP HP is predicted to form on the genomic and/or antigenomic RNA level (nt 56-78, Fig. 1B) and is thought to be relevant to regulation of transcription by VP30 (Weik et al., 2002; Schlereth et al., 2016; Xu et al., 2017). In the mRNA-enriched library, most NP mRNA reads had their 5'-end at the expected position 56 and a minor fraction at nt 57 and 58 (Fig. 2B). This implies that the viral polymerase can also initiate transcription at positions 57/58, although at present we cannot exclude that all transcripts start with G<sub>56</sub>, but lost this or the following nucleotide during library construction, sequencing or data processing.

## Northern Blot and qRT-PCR confirm the presence of *leader*RNAs in EBOV-infected cells

To validate our finding of abortive *leader*RNAs in EBOV-infected cells by RNA-Seq (Fig. 2A), we performed Northern Blot analyses with a leader-specific T7-transcribed and digoxigenin-labeled probe that exhibits full complementarity to nt 1-80 of the expected leader transcript (Fig. 3A, left). Indeed, we were able to detect *leader*RNAs by using 20-40 µg of total RNA extracted from EBOV-infected HuH7 cells. *In vitro*-transcribed (ivt) RNAs identical to antigenome nt 2-78, 56-158 as well as an ivt RNA mimicking genome nt -1 to -154 (Fig. 1B) were used to confirm probe specificity. The major fraction of leader transcripts apparently had a length of < 78 nt. Heterogeneous transcript lengths are consistent with the RNA-Seq data (Fig. 2A).

In a third approach we used quantitative real-time PCR (qRT-PCR) to estimate the ratio of leader transcripts and the first NP mRNA. For this purpose, we designed three primer sets for amplification of NP mRNA, *leader*RNA and longer antigenomic RNA as a measure of replicative RNA (cRNA; Fig. 3B). Notably, only two of the three primer sets are specific for a single viral RNA species (NP mRNA and cRNA), whereas the leader primer set amplified *leader*RNA as well as cRNA. Using this qRT-PCR setup, we again confirmed leader transcript synthesis as the primer set detecting both, *leader*RNA and cRNA, resulted in lower  $C_T$  values of ~ 16 than the cRNA-specific primer set ( $C_T$  ~ 18) in all samples analyzed (exemplary experiment shown in Fig. 3C).

### *Leader*RNA amounts are ~ 9-fold lower than that of the NP mRNA

$C_T$  values obtained with the NP mRNA-specific primer pair were clearly lower than those obtained with the *leader*RNA+cRNA primer pair ( $C_T$  ~ 13 versus  $C_T$  ~ 16), indicating that NP mRNAs are more abundant than *leader*RNAs. As the *leader*RNA primer set simultaneously detected *leader*RNA and cRNA, we indirectly calculated the ratio of NP mRNA:*leader*RNA (see Fig. S3 for details). A total of three biological replicates were analyzed in technical duplicates to determine an average ratio of ~ 1:9 for *leader*RNA:NP mRNA amounts and ~ 4:1 for *leader*RNA:cRNA amounts (Fig. 3D and Fig. S3).

### *Leader*RNAs are also produced in MG systems

We also performed RNA-Seq using small RNA (< 200 nt) preparations derived from cells transfected with monocistronic EBOV MGs. This revealed the presence of abortive *leader*RNAs as well (Fig. S4A), with most 3'-ends also mapping to sequences between PE1 and PE2. Here, ~ 65 % of the reads had their 3'-end between position 70 and 79



and some reads (5.7 %) were mapped to the region of nt 80-89, suggesting that the *leader*RNA pattern may be shifted to slightly longer *leader*RNAs in MG-transfected HEK293 cells versus EBOV-infected HuH7 cells. Nonetheless, our sequencing data demonstrate that the synthesis of abortive *leader*RNAs is not restricted to virus infections but can be investigated using EBOV-specific MG systems as well, thereby enabling mechanistical and biochemical studies outside BSL-4 laboratories.

We then devised a qRT-PCR approach for leader transcript detection and quantification in MG-derived RNA samples (Fig. S2; Fig. 4A), somewhat different from the qRT-PCR approach applied to virus-infected cells (Fig. 3B). Instead of a specific mRNA primer set annealing to the NP mRNA's very 3'-end, including its poly(A) tail, we now used a well-established primer pair for the Renilla luciferase reporter gene that was demonstrated to enable reliable quantification of mRNA and cRNA (Hoenen et al., 2010; Bach et al., 2019a). The primers for *leader*RNA+cRNA as well as cRNA amplification were the same as in Fig. 3B. For qPCR of cRNA, however, we used the same primer pair as for *leader*RNA+cRNA amplification (primers RT\_leader+cRNA and qPCR\_leader+cRNA), resulting in the same PCR product length of 66 nt and thus equalizing primer efficiencies for cRNA and *leader*RNA+cRNA (cf. Fig. 3B and S2). Furthermore, RT-qPCR of Firefly luciferase mRNA, synthesized from co-transfected plasmids, served as an internal standard to cancel out fluctuations in plasmid transfection efficiency (Materials and Methods; Bach et al., 2019a). Using this qRT-PCR setup we quantified *leader*RNA levels and calculated  $2^{-\Delta CT}$  values. As both primer sets used for Renilla mRNA or *leader*RNA amplification simultaneously amplify cRNA and primer efficiencies were comparable (both had E values of  $\sim 2.00$ ), we could directly calculate an mRNA:*leader*RNA ratio of  $\sim 68:1$  by dividing the  $2^{-\Delta CT}$  values for *leader*RNA+cRNA and mRNA+cRNA. Hence, in the MG context, mRNA levels largely exceed those of *leader*RNA and also cRNA (Fig. 4A).

### **Viral *leader*RNA synthesis is specifically initiated opposite of genome nt -2**

A recent study reported that antigenome synthesis is not initiated opposite to the 3'-terminal G, but opposite to C<sub>-2</sub> (Deflubé et al., 2019; illustrated in Fig. 1B). This raised the question whether the viral polymerase initiates *leader*RNA opposite to C<sub>-2</sub> as well or at another site as in the RSV system (see above). Hence, we took a closer look at *leader*RNAs in our RNA-Seq libraries (Fig. 2A and Fig. 4B) and could confirm that *leader*RNA reads indeed lack the nucleotide complementary to genome pos. -1 in RNA from EBOV-infected and MG-transfected cells (Fig. 4B, left two columns). To ascertain that the absence of the first nucleotide is not an artefact of the RNA-Seq procedure or

data processing, we performed an RNA-Seq experiment under identical conditions using a small RNA (< 200 nt) preparation from HEK293 cells spiked with a synthetic 65-meric *leader*RNA mimic (carrying a 5'-triphosphate; for sequence information see Table S2) that included nt C<sub>1</sub>. Sequencing revealed that almost 100 % of the *spike-in* RNA reads included the 5'-terminal C<sub>1</sub> nucleotide (Fig. 4B, right column), while it was lacking in RNA samples derived from either EBOV-infected or MG-transfected cells. This indicates that not only viral antigenomes (cRNAs) but also *leader*RNAs are initiated opposite genome position -2.

### **First genome nucleotide is dispensable for replication but preferred in transcription**

As the results by Deflubé et al. (2019) and our data (Fig. 4B) suggest genome position -2 as the starting nucleotide for viral polymerization, we addressed the question whether the genome's first nucleotide is not only dispensable for replication but for mRNA transcription as well, considering that *leader*RNAs might be pre-products of viral NP mRNA synthesis (Fearn and Plemper, 2017). We therefore constructed minigenomes lacking the 3'-terminal first (G<sub>-1</sub>; variant Δ1) or the first two (G<sub>-1</sub> and C<sub>-2</sub>; variant Δ2) nucleotides, as well as a variant carrying an additional 3'-nucleotide (variant +G, Fig. 5A). We first measured reporter activity of these variants in the context of a replication-competent minigenome backbone (RC MG). The Δ1 variant was even slightly more efficient than the wt construct (Fig. 5B). Deletion of nucleotides 1 and 2 (variant Δ2), however, strongly decreased reporter activity. Remarkably, an additional G residue at the genome 3'-end (variant G+) substantially increased reporter activity relative to the wt MG (Fig. 5B). We then analyzed the same genome variants as part of the replication-deficient MG backbone that restricts the viral polymerase to transcription. Strikingly, all three mutant variants showed a clear defect in this setup (Fig. 5C), indicating that the wt 3'-end is optimal for viral transcription; even the +G variant, most active in the RC MG system, showed decreased transcriptional activity in the RD MG context. Variant Δ2 was essentially reduced to background (-L) levels, supporting the key role of C<sub>-2</sub> for initiation of EBOV transcription and replication (Deflubé et al., 2019). The levels of mRNA, cRNA and vRNA, derived from the same RC MG-transfected cells as used for reporter gene assays, showed the same ranking of the tested variants as observed in the reporter assay (cf. Fig. 5D-F and 5B).

### **Abortive *leader*RNA synthesis does not result from alternative GE signals**

The RNA-Seq results indicate that leader transcripts are predominantly terminated immediately after the polymerase has passed the NP GS signal in EBOV-infected (Fig. 2A) or MG-transfected cells (Fig. S4). For RSV, it was noticed that leader promoter nucleotides which were shown to be essential for both transcription and replication are also conserved at the corresponding positions in the GS of the L gene (Noton and Fearn, 2015). This suggested the possibility that *leader*RNAs (~ 25 nt in the RSV system) may be terminated by an unidentified termination (GE) signal, thus also employing a mechanism similar to the start-stop mechanism ubiquitously used by NNS viruses in mRNA transcription. To address the possible presence of an alternative GE signal that could explain the production of abortive *leader*RNAs in EBOV, we started from a bicistronic minigenome mimicking the first two EBOV genes, which had the following features (Fig. 6A, top): it comprised the authentic 3'-leader and 5'-trailer sequences, the native NP-VP35 gene border, but the protein-coding regions of the NP and VP35 genes replaced with two different luciferase genes; in addition, the L 3'-UTR substituted for the VP35 3'-UTR, and the VP35 5'-UTR hairpin was replaced with another NP 5'-UTR hairpin (construct termed bici "wt NP-NP"). To find out if the NP hairpin itself or if *cis*-acting sequences within the NP hairpin structure are sufficient for alternative transcript termination, we first mutated the NP GE signal at the NP-VP35 gene border (Fig. 6A, bici NPGEmut.-NP, second from top). In the third construct, a derivative of the second, we exchanged the NP hairpin at the 2<sup>nd</sup> cistron with a sequence analogous to genome nucleotides -1 to -78, including the full-length leader as well as the NP hairpin, to investigate if either sequences between the leader and NP hairpin are recognized as alternative GE signals or if the leader and NP hairpins may form long-range interactions that induce transcript termination (Fig. 6A, bici NPGEmut.-leNP). Note that we only expected synthesis of mRNA 2 and thus Firefly luciferase (Fluc) reporter gene activity if an alternative GE signal compensates the defective NP GE signal at the NP-VP35 gene border. As positive control we included an additional derivative construct in which we replaced the NP hairpin at the 2<sup>nd</sup> cistron with the VP40 5'-UTR hairpin (see Fig. 1A; Fig. 6A, bottom). The VP40 HP harbors an overlapping GS-GE signal, thus in our bicistronic minigenome context it introduces a functional GE signal that enables termination of mRNA 1 despite a defect GE signal at the NP-VP35 gene border (Fig. 6A, bici NPGEmut.-VP40). In luciferase reporter assays, we did not observe substantial Fluc reporter activity for minigenome constructs 2 (NPGEmut.-NP) and 3 (NPGEmut.-leNP) that comprise the mutant NP GE signal, while our positive control resulted in Fluc reporter activity comparable and even higher than that of the wt NP-NP MG construct (Fig. 6B-C). Hence, it is unlikely that an alternative GE signal within the NP hairpin

sequence and/or leader sequence mediates transcript termination. Residual reporter gene levels above the background (-L) control in the case of constructs 2 and 3 might be explained by residual induction of termination by the mutated GE signal.

### VP30 suppresses leader transcript synthesis

It was previously reported that the NP hairpin structure regulates the VP30-dependent transcription initiation at the first GS (TSS) signal (Weik et al., 2002). While the native NP hairpin structure enforced VP30-dependent transcription initiation, a destabilized NP hairpin structure (NheI NP) relaxed the strict VP30 requirement for transcription at the TSS. To examine if VP30 affects *leader*RNA synthesis, we applied the qRT-PCR setup illustrated in Fig. S2. As expected, mRNA+cRNA levels, predominantly representing mRNA levels (Fig. 4A), largely decreased in the absence of VP30, while cRNA alone was significantly increased (Fig. 7A, B). This is in line with previous qRT-PCR data demonstrating that a decrease of transcription correlates with an increase of replication (Kruse et al., 2018, Bach et al., 2019a). In addition, *leader*RNA+cRNA levels were also increased in the absence of VP30 (Fig. 7C). This increase was even more pronounced (~ 1.6-fold; Fig. 7C) compared to the increase observed for cRNA alone (~ 1.2-fold; Fig. 7B), emphasizing that *leader*RNA levels increase as well in the absence of VP30. A comparison of  $2^{-\Delta CT}$  values obtained by the *leader*RNA+cRNA and the cRNA-specific primer sets illustrates this more distinctly (Fig. 7D), revealing that the increase in *leader*RNA+cRNA levels in the absence of VP30 exceeds that of cRNA alone. Our findings thus indicate that leader transcript synthesis is reduced in the presence of VP30.

RNA-Seq, as introduced in Fig. 2, was performed with RNA from cells transfected with the wt NP MG, a MG construct harboring a destabilized hairpin (NheI NP) and variant  $\Delta 5'$  spacer lacking the capacity to form a hairpin structure owing to a 12-nt deletion in the spacer between PE1 and PE2 (Fig. 8A). In line with the qRT-PCR results, we observed an increase in *leader*RNA amounts in the absence of VP30 for the wt NP MG (Fig. 8B). The same pertains to the two mutant MGs (Fig. 8B).

Weik et al. (2002) previously inferred from Northern blot results that the NheI NP mutation even enabled substantial mRNA synthesis in the absence of VP30, which led the authors to suggest that transcription became largely VP30-independent. We thus compared the NP mRNA levels in our RNA-Seq data for the NheI NP and  $\Delta 5'$  spacer mutants with those of the wt NP MG (Fig. 8C). We indeed observed a substantial increase for the NheI NP as well as a mild increase for the  $\Delta 5'$  spacer mutant in mRNA reads compared to the wt MG in the absence of VP30. However, our data also indicate

that transcription, although somewhat relaxed in the absence of VP30, remains largely VP30-dependent even for the NheI NP and  $\Delta 5'$  spacer mutant constructs with weakened or abolished secondary structure potential. The  $\Delta 5'$  spacer mutant showed both reduced *leader*RNA levels and mRNA levels (Fig. 8B, C) in the presence of VP30. It appears that the 12-nt deletion in the  $\Delta 5'$  spacer mutant moderately reduces viral polymerase activity in general.

### **Site-specific transcription initiation can occur in the absence of VP30**

As we could observe VP30-independent mRNA synthesis initiation at the TSS both for the NheI NP and  $\Delta 5'$  spacer MG variants (Fig. 8C), we wondered if deviations from the native NP hairpin structure might affect selection of the transcription initiation site. However, RNA-Seq revealed that, even in the absence of VP30, mRNA synthesis was initiated opposite genome nt -56 in case of the  $\Delta 5'$  spacer MG (Fig. 8D, E), the NheI NP MG (Fig. S5) and the wt NP MG (Fig. S5). Thus, recognition of the TSS and site-specific initiation of mRNA synthesis is basically the same in the presence or absence of VP30. It has yet to be noted that read numbers were very low in the absence versus presence of VP30, particularly for the wt NP MG, in line with the crucial role of VP30 for the efficiency of viral transcription.

### ***Leader*RNA termination is independent of RNA structure formation potential at the TSS**

Overall, 3'-ends of *leader*RNAs, as inferred from RNA-Seq, map to the same region for EBOV-infected cells as well as cells transfected with the wt NP MG and the two aforementioned mutant constructs (Fig. 9A-D; for more details, see Discussion). As this also pertains to the  $\Delta 5'$  spacer MG devoid of any secondary structure, we can conclude that termination of leader transcripts is not influenced by the potential to form RNA structures in the spacer between PE1 and PE2.

### **Initiation and termination of *leader* transcripts is independent of VP30**

We then comparatively analyzed the small RNA libraries for *leader*RNA initiation and termination sites in RNA samples derived from cells transfected with the wt NP MG versus the NheI NP and  $\Delta 5'$  spacer MGs, either in the presence or absence of VP30. For all three MG variants and independent of VP30, *leader*RNA 5'-ends mapped to position 2 (Fig. 1B, bottom; Fig S6). Only ~ 0.2 % of reads mapped to position 1, similar to *leader*RNAs isolated from EBOV-infected cells (Fig. 4B). In all cases, the RNA-Seq data provided no evidence for substantial changes in the respective 3'-end patterns of

*leader*RNAs upon omission of VP30 (Fig. S4). We conclude that VP30 neither affects the position of *leader*RNA initiation nor the pattern of termination sites.

### **Leader transcripts are not synthesized if hexamer spacing between PE1 and PE2 is violated**

The NP hairpin represents the major part of the spacer region separating PE1 and PE2 of the leader promoter (Fig. 1B). It was recently shown that hexamer phasing from PE1 to PE2 (termed “rule of 6”) is not only crucial for replication but also EBOV transcription initiation in the 3'-leader (Bach et al., 2019a). Furthermore, it could be shown that nucleotides -54 and -55 in PE1 are part of both the replication promoter as well as the putative transcription promoter, thereby suggesting at least partial overlap of signals relevant to replication and transcription. Furthermore it strengthens the notion that the viral polymerase enters the genome template from the very 3' end for both, replication and transcription. This raises the question whether hexamer phasing between PE1 and PE2 affects *leader*RNA synthesis as well. To address this issue, we used two MG variants that violate the “rule of 6”, one carrying the NP hairpin with a 1-nt deletion near the stem and a variant in which the NP hairpin was replaced with the corresponding element of the VP35 gene (Fig. 10A). In contrast to the wt NP hairpin with a hexameric spacer of 30 nt (genome pos. -51 to -80), both mutant minigenomes comprised only 29 nt in this region. As shown by qRT-PCR, the two mutant minigenomes were deficient in antigenome (cRNA), mRNA and *leader*RNA synthesis (Fig. 10B-D). We conclude that hexamer phasing is generally essential for initiation of RNA synthesis in the 3'-leader promoter.

## **DISCUSSION**

In line with the “single-entry and start-stop” transcription model, many NNS viruses were shown to synthesize short transcripts antisense to the genome 3'-leader sequence (termed leader transcripts or *leader*RNA) before transcription initiation at the first GS signal (Fearn and Plemper, 2017; Honda et al., 2017; Noton and Fearn, 2015; Ogino and Green, 2019). Our findings combined with those of others (Deflubé et al., 2019) indicate that *leader*RNAs and cRNAs are both initiated at the penultimate nucleotide of the EBOV genome. This feature of a single initiation site at the genome's 3'-end is shared with other NNS viruses such as VSV (*Rhabdoviridae*) and SeV (*Paramyxoviridae*), which yet initiate *leader*RNAs and cRNAs at the genome's very 3'-end (Chuang and Perrault, 1997; Horikami and Moyer, 1991). For VSV and SeV, a mechanism has been proposed according to which the viral polymerase first synthesizes a leader transcript that is only elongated to cRNA when immediately encapsidated by the

nucleoprotein (NP). At low NP concentrations early in infection, the leader transcript is terminated and the polymerase initiates transcription at the first gene and becomes programmed to recognize the GS/GE signals and to cap and polyadenylate the individual mRNA transcripts (Chuang and Perrault, 1997; reviewed in: Fearn and Plemper, 2017; Noton and Fearn, 2015; Whelan et al., 2004). The model is supported by studies using transcription-competent VSV cores or purified SeV preparations (Chuang and Perrault, 1997; Vidal and Kolakofsky, 1989), where the authors reported roughly equimolar amounts of leader and NP mRNA, respectively. However, a VSV mutant strain (polR1) produced about twofold more NP mRNA than *leader*RNA, suggesting the possibility that the viral polymerase complex might have the capacity to also initiate RNA synthesis directly at the TSS of the NP gene (Chuang and Perrault, 1997).

In the present study we analyzed *leader*RNA levels in EBOV-infected and MG-transfected systems (Fig. 3, 4). Estimation of the mRNA:*leader*RNA ratio revealed an excess of mRNA levels of 9:1 (NP mRNA:*leader*RNA) during virus infection and 68:1 (Renilla mRNA:*leader*RNA) in MG-transfected cells (after 24 h and 48 h respectively). Considering that the MG setup provides constant excess amounts of NP molecules, this finding appears to be consistent with the abovementioned model predicting that termination of *leader*RNA synthesis is suppressed at high NP levels. The model would imply that most *leader*RNAs are elongated to cRNAs, thus boosting replication in the MG system. However, the MG system yields almost 100-fold higher amounts of mRNA than cRNA (Bach et al., 2019a), suggesting that NP availability per se is not the primary cause of low *leader*RNA levels in the MG system. The excess of mRNA over *leader*RNA in general, and differences in *leader*RNA levels in viral infection versus MG transfection, may also reflect differential increases in the decay rate of *leader*RNAs. Indeed, for other NNS viruses, lower *leader*RNA levels in cells versus *in vitro*-transcription assays and different levels in different cell types were observed (Horikami and Moyer, 1991; Vidal and Kolakofsky, 1989). Yet, there is also the example of *leader*RNAs detected in equimolar amounts to the first mRNA in HEK293 cells infected with Borna disease virus (Honda et al., 2017). This suggests that *leader*RNAs may not necessarily be rapidly degraded in this cell line that we also used for our MG system. Thus, at present it remains an open question to which extent the measured ratios of NP or Renilla mRNA:*leader*RNA levels include contributions from differential decay rates of the two RNA species.

The qRT-PCR data in Fig. 7D indicate that *leader*RNA levels in the MG system are similar to, though somewhat higher than cRNA levels. Both levels increase in the absence of VP30 (Fig. 7D), conditions where mRNA synthesis is shut down (Fig. 7A; Fig.

8C, wt NP MG). This indicates that *leader*RNAs and cRNAs are products of the viral polymerase acting in its replication mode. This notion is also supported by the RNA-Seq data revealing higher leader transcript amounts in the absence of VP30 (Fig. 8B). SeV is a paramyxovirus that has its TSS at the identical position as EBOV (nt -56). For SeV, MGs with an additional (external) promoter fused to the native genome 3'-end were constructed, which moved the TSS (GS1; now part of the internal promoter) 90 nt more inwards (from pos. -56 to -146). Viral transcription still occurred (~ 40 % activity), even if PE1 sequences at the GS1 of the internal promoter, which are essential for replication, were deleted (Vulli  moz et al., 2005). The results indicated that the viral polymerase generally accesses the nucleocapsid genome from the 3'-end, consistent with the model that polymerase begins to synthesize a cRNA or abortive *leader*RNA; in case of *leader*RNA termination, the enzyme appears to be able to scan the template RNA backwards or forward for a GS signal, or restarts synthesis at the genome 3'-end if no functional GS signal is recognized. For EBOV, higher leader transcript amounts in the absence of VP30 would then indicate that the polymerase is inefficient without VP30 in recognizing the TSS and thus more frequently reinitiates RNA synthesis at the genome 3'-end. It is interesting to note that EBOV *leader*RNAs are terminated after the RdRp has passed the TSS, meaning that the enzyme has to move backwards on the genome for transcription initiation as well as reinitiation at the genome 3'-end. The capacity of EBOV RdRp to move backwards or forward to the next GS signal after terminating at a GE signal, was demonstrated in a previous study (Brauburger et al., 2016). However, a clear difference between *leader*RNA termination and termination at internal GE signals is the absence of a clear sequence signature for termination of *leader*RNAs (see discussion below), consistent with the model that *leader*RNAs are terminated because of insufficient encapsidation.

A remaining key question is whether each transcriptional initiation event requires the synthesis of an abortive replication product (= *leader*RNA). Such a mechanism would predict equimolar synthesis of *leader*RNAs and mRNA 1 (as observed for VSV and SeV, see above). Another possibility is that the RdRp enters the genome at the very 3' end and is able to undergo (supported by VP30) a conformational rearrangement (without polymerization) to position the TSS in its active site to initiate mRNA synthesis. Such a scenario would imply that the polymerase gains access to replication and transcription promoter elements via binding to the genome 3'-end. We previously showed that adherence to the "rule of 6" in the spacer region between PE1 and PE2 is crucial for both replication and transcription, and that nt -54/55 in PE1 are essential for transcription (Bach et al., 2019a). Thus, sequence elements upstream and downstream of the transcription start sequence are crucial for replication and transcription. Likewise, the



present study suggests that, for EBOV transcription initiation, the viral polymerase needs to contact the very 3'-end directly (presumably as part of the transcription promoter). The genome variant  $\Delta 1$  lacking the  $G_{-1}$  residue showed a ~ fourfold decrease in reporter activity relative to the wt genome (Fig. 5C) in the context of a replication-deficient MG backbone. In contrast, variant  $\Delta 1$  was as active as the wt construct in the replication-competent MG system (Fig. 5B, D-F) and retained ~ 70 % of replication activity in another study (Deflubé et al., 2019). This observation may be explained as follows: as long as viral replication is enabled, the 3'-terminal residue (nt -1) can be added to newly synthesized vRNAs by the viral polymerase, possibly in a terminal transferase reaction (Deflubé et al., 2019). Furthermore, only low amounts of functional nucleocapsids are thought to be assembled from minigenome vRNAs produced by non-viral RNA polymerases (such as T7 RNAP) and NP protein in the absence of viral replication (Blakqori et al., 2003), explaining the low levels of reporter activity in RD MGs (Hoenen et al., 2010). Thus, viral replication in RC MGs will not only generate vRNAs carrying the  $G_{-1}$  residue, but those vRNAs will also be part of more functional nucleocapsids and thus outcompete the pre-synthesized T7 vRNAs as RNA templates for viral mRNA synthesis. This is expected to have masked those transcriptional defects that were revealed in the RD MG system (Fig. 5C). Hence, the RD MG system appears better suited to analyze the effect of genome 3'-end variants on transcription. It should be mentioned that RD MGs, although giving rise to rather low reporter activities, provide biologically relevant readouts that are fully in accordance with results obtained with RC MGs, as shown for the dependence of viral transcription on VP30 and the "rule of 6" (Hoenen et al., 2010; Bach et al., 2019a).

The +G variant was also more active than the wt construct in the RC MG system, and here most evident on the cRNA level (Fig. 5B, D-F). This variant had only ~ 40 % activity in the RD MG system (Fig. 5C). In summary, our findings demonstrate that the EBOV polymerase strongly prefers genome 3'- $G_{-1}CC$  ends in transcription, whereas replication is largely permissive to the absence of  $G_{-1}$  and seems even be enhanced by adding another G residue to the 3'-end. It will be interesting to see if genomes packaged into infective EBOV particles carry 3'- $G_{-1}CC$  ends to support primary transcription during early infection.

The aforementioned findings suggest mechanistic differences in initiation modes of replication and transcription, possibly an intimate polymerase contact to nt  $G_{-1}$  in the transcription mode. The 2<sup>nd</sup> genome nt ( $C_{-2}$ ), however, is essentially indispensable for transcription and cRNA synthesis (Deflubé et al., 2019; Fig. 5). It may be part of the

replication and transcription promoter or it may be crucial for NP phasing. For paramyxoviruses, the only other family of NNS viruses that utilizes bipartite replication promoters (Murphy and Parks, 1999; Tapparel et al., 1998), a proper spacing and/or NP phasing of the genome 3'-end and the TSS was inferred to be important to transcription. In the paramyxoviruses SeV and HPIV3, transcription still occurred when most of the leader and/or PE2 were deleted, provided that a certain NP phase at the TSS was maintained (Cordey and Roux, 2006, 2007; Hoffman and Banerjee, 2000b; Vulli  moz et al., 2005). In a recent study we showed that EBOV transcription was eliminated by a dinucleotide deletion (nt -54 and -55) directly 3' of the TSS. Thus, these 2 nt are either part of the transcription promoter or might have changed NP phasing of the leader by positioning the TSS closer to the genome 3'-end (Bach et al., 2019a). We have demonstrated (this study; Bach et al., 2019a) that hexamer phasing in the PE1-PE2 spacer is obligatory for mRNA, cRNA and *leader*RNA synthesis. We recently proposed a model according to which productive binding of the EBOV polymerase to the 3'-terminal region of the genome, as a prerequisite for any polymerization activity, requires that NP molecules are assembled in the correct register in the region of nt -51 to -80 and further inwards (Bach et al., 2019a). Based on this model, termination of *leader*RNAs in the region of nt ~ 60-80 may then be caused by inefficient removal of NP from the template RNA in this region or inefficient NP encapsidation of nascent *leader*RNAs. For viral transcription, NP molecules have to be removed from the TSS (nt -56 to 67), and this may be achieved with the help of VP30 and/or VP35. (Cox et al., 2014; Shu et al., 2019).

We analyzed potential determinants of *leader*RNA termination near the TSS in the region of the NP hairpin structure. One possibility we considered is termination caused by NP hairpin formation (antigenomic nt -56 to -78, Fig. 1B). However, the majority of *leader*RNAs not only terminated after ~ 60 to 80 nt in EBOV-infected cells and those transfected with the wt NP MG, but also in cells transfected with the NheI NP and the  $\Delta$ 5' spacer MGs (Fig. 9A-D). As the  $\Delta$ 5' spacer MG lacks any structure formation potential in the region between PE1 and PE2 (Fig. 8A), we can exclude a key role of hairpin formation in *leader*RNA termination. It is noteworthy that the largest fraction of *leader*RNA reads > 80 nt was observed for RNAs derived from cells transfected with the NheI NP MG (Fig. 9C), indicating that this mutant MG may be phenotypically unique and thus not representative, for reasons not yet understood. The lowest fraction of reads > 80 nt was seen in libraries of small RNAs derived from EBOV-infected cells.

For VSV, mutant N proteins were reported to suppress *leader*RNA termination and to favor *leader*RNA readthrough (Perrault et al., 1983), suggesting a key role of NP in *leader*RNA termination. It is thus not too farfetched to consider the possibility that binding

of NP molecules to the UN<sub>5</sub> hexamers in PE2 of EBOV may provide a barrier that induces termination of *leader*RNAs in the region preceding PE2. However, this can be excluded as well since *leader*RNAs derived from the  $\Delta 5'$  spacer MG cover the same size range as those derived from the wt NP MG and thus predominantly terminate within PE2 (Fig. 9D).

In *Pneumoviridae*, such as in the Respiratory Syncytial Virus (RSV), leader transcripts terminate at a U-rich sequence directly preceding (3' of) the TSS. Such sequence elements were previously shown to be important for transcription in some of the NNS viruses (Cowton et al., 2006; Hinzman et al., 2002; Li and Pattnaik, 1999; Whelan and Wertz, 1999). As GE signals contain U stretches as well, U-rich sequences directly preceding the TSS were considered as alternative leader termination signals. However, in contrast to mRNA transcripts, *leader*RNAs are not precisely terminated at a distinct length but show heterogeneous lengths, suggesting that the polymerase is in a rather unprocessive mode during their synthesis (Ghosh et al., 1996; Horikami and Moyer, 1991; Kurilla et al., 1985; Leppert et al., 1979; Tremaglio et al., 2013; Vidal and Kolakofsky, 1989). Moreover, such a mechanism is even more unlikely for EBOV *leader*RNAs because the EBOV leader lacks such continuous U stretches 3' of the TSS. Furthermore, we were not able to experimentally detect any strong alternative GE signals within the EBOV leader-NP HP region (Fig. 6) that might explain *leader*RNA termination. Thus, we rule out that a sequential stop-start transcription mechanism, as operational at internal EBOV genes, also applies to the EBOV leader-NP junction. Similar conclusions were drawn for the VSV leader-N gene junction (Whelan et al., 2004). In conclusion, the EBOV polymerase complex synthesizes *leader*RNAs of heterogeneous length (~ 60 to 80 nt), largely independent of hairpin structure formation potential at the TSS and little affected by shortening the distance of PE2 to the genome 3'-end by 12 nt. The fact that the *leader*RNA length spectrum is also ~ 60 to 80 nt for the  $\Delta 5'$  spacer MG makes it unlikely that specific sequences in the spacer between PE1 and PE2 induce termination of *leader* transcripts. Also, omission of VP30 did not change the length spectrum of *leader*RNAs and tended to enhance the amounts of *leader*RNAs (Fig. S4, 7D, 8B). A straightforward interpretation of our findings is that the VP35/L polymerase complex synthesizes *leader*RNAs without VP30, and the latter only associates with the polymerase complex during the process of transcriptional initiation at the TSS.

Noteworthy, our testing for alternative leader GE signals in bicistronic minigenomes revealed that a small amount of mRNA2 (~ 6 %) was produced (Fig. 6C) upon disruption of the native intergenic GE signal and upon replacing the intergenic VP35 hairpin with

the NP hairpin structure or the 3'-terminal 78 nt of genome (Fig. 6A, 2<sup>nd</sup> and 3<sup>rd</sup> construct). Apart from the possibility that the mutated GE signal retained some residual termination activity, the findings might alternatively suggest that either the RdRp is capable of initiating mRNA synthesis directly at internal GS signals with low efficiency or that the presence of the NP hairpin structure itself is able to stall the polymerase complex which then allows some of the polymerases to scan the template for the GS signal and reinitiate mRNA synthesis. However, this is speculative and, in any case, would represent a minor low-efficiency pathway.

The exact function of EBOV leader transcripts in the context of replication and transcription still remains to be defined. *Leader* RNAs might be sheer regulatory by-products of replication, e.g. resulting from inefficient NP encapsidation of nascent cRNAs and triggering viral transcription as proposed for VSV (Blumberg et al., 1981). A key role for NP in *leader* RNA termination was inferred from the identification of NP mutants of VSV that increase *leader* RNA readthrough (Perrault et al., 1983). Likewise, influenza virus short RNAs corresponding to the 5' end of the genome leader regions have been suggested to arise simultaneously to vRNA synthesis and to regulate the switch from viral transcription to replication (Perez et al., 2010; Umbach et al., 2010). Leader transcripts in other NNS viruses were suggested to be recognized by RIG-I, thus inducing an innate immune response (Louber et al., 2014; Oh et al., 2016; Plumet et al., 2007). Conversely, there are several examples of short leader/trailer RNAs subverting RIG-I mediated immune responses as a potential escape mechanism. In RSV infection, *leader* RNAs (LeC-containing RNAs) were suggested to revoke RIG-I recognition by binding to the cellular La autoantigen (Bitko et al., 2008). Trailer RNA transcripts of SeV were reported to bind to the cellular RNA binding protein TIAR and to exert anti-apoptotic effects and to avert the cellular stress granule response (Iseni et al., 2002). For the RSV trailer RNA, an involvement in subverting stress granule responses was proposed as well (Hanley et al., 2010). Considering that EBOV *leader* RNAs are initiated opposite to genome position -2, it is conceivable that they anneal to non-encapsidated genome 3'-ends during RNA synthesis, generating a duplex with a 1-nt 3'-overhang that is unfavorable for RIG-I activation (Schlee et al., 2009). The only other virus initiating RNA synthesis at position 2 is the Tacaribe virus, a segmented negative strand RNA virus of the arenavirus family. In a proposed prime-realign mechanism, the internally (at pos. +2) initiated RNA dinucleotide 5'-pppGpC slips backwards on the template RNA to generate a 1-nt 5'-overhang which prevents RIG-I recognition (Garcin and Kolakofsky, 1992; Marq et al., 2010). EBOV leader transcripts (*in vitro* T7 transcripts) showed some RIG-I activation, but it remained unclear whether side products generated by T7 transcription were responsible for this observation (Louber et al., 2014). At least RNA secondary

structure prediction and RNA probing in solution indicates that *leader*RNAs fold into a hairpin with a three-nt 5'-overhang unfavorable for RIG-I recognition (Fig. 1B). It is also a possibility that *leader*RNAs may act as baits to sequester RIG-I, thereby dampening RIG-I-mediated immune responses. We have so far tested a *leader*RNA 65-mer (including nt G<sub>-1</sub>), but have neither seen significant RIG-I activation nor considerable attenuation of RIG-I activation by VSV RNA transfected together with the EBOV *leader*RNA 65-mer (unpublished results).

## MATERIAL AND METHODS

### *Cell culture*

Human embryo kidney cells (HEK293; ATCC®, DMSZ ACC 305) and human hepatocellular carcinoma cells (HuH7; ATCC®, JCRB 0403) were cultivated at 37 °C and 5 % CO<sub>2</sub> in a humidified atmosphere in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 50 U/ml penicillin (+P), 50 µg/ml streptomycin (+S), 2 mM L-glutamine (+Q) and 10 % fetal calf serum (FCS). All components were purchased from Thermo Fisher Scientific (TFS). Cloning and propagation of MG plasmids was conducted in the *E. coli* DH5α strain using standard microbiological procedures.

### *Plasmids*

Plasmids used in MG assays, such as derivatives of plasmid pCAGGS encoding the Zaire EBOV nucleocapsid proteins VP30, NP, VP35 and L, as well as plasmids coding for the T7 RNA polymerase, the EBOV-specific wild-type MG (pANDY 3E5E) or a wild-type replication-deficient variant thereof, were described earlier (Hoenen et al., 2006, 2010). For the construction of bicistronic minigenomes plasmids pA-3E5E-GLuc-GFP and pCAGGS\_Luc2 were used. For details on pA-3E5E-GLuc-GFP see Supplementary Table S1. Construction of plasmid pCAGGS\_Luc2 has been described earlier (Biedenkopf and Hoenen, 2017).

### *Cloning of replication-competent (RC) and replication-deficient (RD) monocistronic MG variants*

3'-leader monocistronic MG variants were derived from the replication-competent (RC) EBOV minigenome pANDY 3E5E. Mutant derivatives of the native leader 3'-terminus were generated by either deletion of the first 3'-G nucleotide (p3E5E\_3'-Δ1), deletion of the 3'-terminal GC dinucleotide (p3E5E\_3'-Δ2) or elongated by an additional G residue (p3E5E\_3'+1) using *DpnI*-based site-directed mutagenesis techniques (see Fig. S1).

Cloning primers are specified in the Supplementary Table S1. Corresponding RD MGs were based on a pANDY 3E5E derivative lacking the terminal 55 nucleotides of the antigenome 3'-trailer replication promoter (Hoenen et al., 2010) using the same primers and mutagenesis setup as described for the RC MG 3'-leader variants. Variants of the wt pANDY 3E5E MG comprising a destabilized NP hairpin structure at the first NP GS signal (p3E5E\_NheI NP) or lacking 12 spacing nucleotides which separate the NP GS signal from replication PE2 (p3E5E\_Δ5' spacer) were described previously (Bach et al., 2019b; Weik et al., 2002). This also pertains to monocistronic RC MGs harboring spacer sequences deviating from the "rule of 6" [p3E5E\_VP35; p3E5E\_NP-1 (stem)] (Bach et al., 2019a, b; Weik et al., 2002). Nucleotide numbers correspond to the EBOV Mayinga (Zaire, 1976) sequence, GenBank accession no. AF086833. Sequence identity of all constructs was confirmed by DNA sequencing.

#### *Cloning of RC bicistronic MG variants*

RC bicistronic MGs used in this study are based on a wild-type RC bicistronic (bici) MG harboring the first two EBOV cistrons, but the NP and VP35 ORFs replaced with the Renilla (RLuc) and Firefly (FLuc/Luc2) luciferase ORF, respectively. The two ORFs, separated by the authentic NP-VP35 intergenic region (IGR), are flanked by the terminal EBOV leader and trailer regions, the NP 5'- and 3'-UTRs as well as the VP35 5'-UTR and the L 3'-UTR (Fig. 6A). Cloning of the wt bici MG and mutant variants thereof was conducted by standard restriction cloning and site-directed mutagenesis, specified in Table S1. Briefly, the wt bici MG (pA-3E5E-RLuc-FLuc) was constructed using a 5-step-cloning strategy based on the plasmids pA-3E5E-GLuc-GFP (used as backbone and harboring the EBOV leader, trailer, UTRs), pCAGGS\_Luc2 (used for exchange of GFP with Fluc) and wt pANDY 3E5E (used for exchange of GLuc with RLuc). Variants comprising an NP 5'-UTR hairpin or VP40 5'-UTR hairpin structure instead of the VP35 5'-UTR hairpin structure at the 2<sup>nd</sup> cistron were cloned by restriction cloning using GeneArt Strings DNA Fragments (TFS). NP GE mutants were cloned by site-directed mutagenesis. Here, the GE signal was inactivated by mutating 3 out of its 6 U residues to adenosines (Fig. 6A) as reported previously (Brauburger et al., 2016).

#### *Infection of HuH7 cells with EBOV*

Work with Zaire EBOV strain Mayinga (Accession number AF086833) was conducted at the biosafety level 4 (BSL-4) laboratory of the Philipps University Marburg. For EBOV RNA preparations,  $8 \times 10^6$  HuH7 cells in DMEM supplemented with 50 U/ml penicillin, 50 µg/ml streptomycin and 2 mM L-glutamine (DMEM<sub>Q+P/S</sub>) were subjected to EBOV at a multiplicity of infection (MOI) of 3 for 1 h at 37 °C. The inoculum was removed and cells

were incubated in DMEM<sub>+3%FCS+Q+P/S</sub> at 37 °C and 5 % CO<sub>2</sub> in a humidified atmosphere. Viral RNA was extracted from EBOV-infected cells 1 day post infection.

#### *EBOV-specific MG assay*

EBOV-specific MG assays were conducted as described earlier (Bach et al., 2019a). Briefly,  $8 \times 10^5$  HEK293 cells per well were seeded in 6-well plates (Greiner) and cultivated at 37 °C and 5 % CO<sub>2</sub> in a humidified atmosphere in DMEM<sub>+10%FCS+Q+P/S</sub> (3 mL/well) 18-24 h prior to transfection. At 60-80 % cell confluency plasmids coding for the T7 promoter-driven EBOV-specific MG variants (250 ng), the EBOV nucleocapsid proteins NP (125 ng), VP35 (125 ng), VP30 (100 ng), L (1000 ng), as well as a plasmid encoding T7 RNA polymerase (250 ng) were transfected into HEK293 cells using TransIT-LT1 (Mirus). In the case of monocistronic MGs, the plasmid pGL4.13 (Promega) encoding firefly luciferase (FLuc) was additionally transfected for normalization of transfection efficiencies. Cells were harvested 48 h post transfection and lysed for luciferase reporter gene assays or RNA extraction.

#### *Luciferase assay*

Luciferase assays were performed as described earlier (Bach et al., 2019a). Briefly, cells were lysed in 200 µL 1× Reaction Lysis Buffer (2x Lysis-Juice; PJK). In case of RC monocistronic MG samples lysates were further diluted 1:50 in ddH<sub>2</sub>O prior to Renilla Luciferase activity measurements. RD monocistronic MG samples and bicistronic RC MG samples were not diluted further. For FLuc activity measurements, undiluted lysates were used for all MG variants. 10 µL of diluted or undiluted lysate were then mixed with either 50 µL of Rluc Reagent (Renilla-Juice Fluid; PJK) or 50 µL of Fluc Reagent (Beetle-Juice; PJK) and bioluminescence, as a measure of reporter gene activity, was detected using a Centro LB 960 luminometer (Berthold Technologies). In case of monocistronic MGs (RC and RD) Renilla luciferase values were normalized to Fluc values to account for differences in transfection efficiency. Results obtained for monocistronic MGs carrying the native wt NP leader were set to 100 %. Relative Renilla or Firefly activities of bicistronic MG samples were determined by normalization to the respective values obtained for the bici NP-NP construct (set to 100 %).

#### *RNA-Seq: cDNA library preparation and Illumina Sequencing*

Total RNA was prepared from EBOV-infected HuH7 cells or from MG-transfected HEK293 cells and either enriched for polyA<sup>+</sup> RNA or small RNAs (< 200 nt). Preparation

of cDNA libraries was performed at vertis Biotechnologie AG in Freising (Germany). For details, see Supplementary Material.

### *Northern blotting*

*Leader*RNA detection by Northern Blotting was performed as described (Hoch et al., 2016) using 20-40 µg of total RNA extracted from EBOV-infected HuH7 cells or 10 ng of *in vitro* transcribed control RNAs (T7 transcripts covering the antigenomic EBOV leader sequences [5' to 3'] 2-78 and 56-158, as well as genomic leader nucleotides -154 to -1; described in Schlereth et al., 2016). RNA samples were adjusted to 1x denaturing loading buffer containing 0.01 % (w/v) bromophenol blue, 0.01 (w/v) xylene cyanol, 1.3 M urea, 33 % (v/v) formamide and 1x TBE and were denatured for 3 min at 95 °C, followed by rapid cooling on ice. Denatured RNAs were separated by 20 % polyacrylamide gel electrophoresis (PAGE) and afterwards transferred onto a positively charged nylon membrane (Roche) using overnight semidry electro-blotting and 0.5x TBE as transfer buffer. RNAs were crosslinked to the nylon membrane using 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC). Crosslinked membranes were transferred to a glass tube and preincubated with ~ 15 ml DIG Easy Hyb™ (Sigma Aldrich/Merck) hybridization solution for 2 h at 68 °C. Afterwards, the pre-hybridization solution was exchanged with ~ 15 ml of hybridization solution containing 5 µL of the digoxigenin-labeled RNA probe (Supplementary Tab. S1/Tab. 38; equivalent to 3'-leader nt 1-78 + GG for enhancing T7 transcription efficiency), followed by incubation overnight at 68 °C. Probes were labeled by *in vitro* T7 transcription using the DIG RNA labeling Kit (Sigma Aldrich/Merck) according to the manufacturer's protocol. After hybridization, the membrane was washed twice with 50 ml stringency buffer 1 (2× SSC, saline sodium citrate; 0.1 % SDS) for 5 min each, followed by 2x washing with 50 mL stringency buffer 2 (0.1× SSC, 0.1% SDS) for 15 min each (20x SSC buffer was purchased from Sigma Aldrich/Merck). RNA was immunologically detected using the DIG northern starter Kit (Sigma Aldrich/Merck) following the manufacturer's recommendations.

### *Extraction and purification of total RNA for qRT-PCR analysis and Northern blotting*

Total RNA of EBOV-infected HuH7 cells was isolated and purified using the mirVana™ miRNA Isolation Kit (with phenol; TFS) according to the manufacturer's protocol. Total RNA of MG-transfected HEK293 cells was isolated using the RNeasy mini kit (QIAGEN) following the manufacturer's protocol. An on-column digestion step using the RNase-Free DNase Set (QIAGEN) for was included during isolation. RNA was eluted in RNase-free water. A 2<sup>nd</sup> DNase-treatment in the presence of 20 U RiboLock RNase Inhibitor (TFS) was performed using Ambion™ DNase I (TFS) at 37 °C for 1 h. Purification was



conducted using Roti®-Phenol/Chloroform/Isoamyl alcohol (Carl Roth). RNA was precipitated by addition of three volumes EtOH:NaOAc [30:1; 3 M NaOAc (pH 5)]. The RNA pellet was washed with 70 % EtOH, air-dried and finally redissolved in RNase-free water.

### *Reverse transcription strategies 1-3*

The strategies for the detection and quantification of the different viral RNAs are schematically illustrated in Fig. 3B (strategy 1) and Fig. S2 (strategy 2). Detection of viral RNAs of MG 3'-end mutants (strategy 3) was conducted as described earlier (Bach et al., 2019a). Briefly, 500 ng (MG system) or 1 µg (EBOV infection) of total RNA were used for reverse transcription (RT) in a total reaction volume of 20 µL using the RevertAid H Minus Reverse Transcriptase (TSF) according to the manufacturer's recommendations. Initial denaturing steps at 65 °C for 5 min (MG samples) or 80 °C for 5 min (EBOV infection samples) were included for denaturation of structured RNAs. For RT-PCR strategy 1 (detection of *leader*RNA+cRNA, cRNA and NP mRNA in total RNA from EBOV-infected cells), the following RT primers were used: RT\_*leader*+cRNA: 5'-ATC TTC CTC ATA GTT ATT CGC ACA CA-3'; RT\_NP mRNA: 5'-TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTC TTA ATT ATA AAA CG-3'; RT\_cRNA s1/s2: 5'-CAA ACC AGG TGT GAT TAC AGT AAC AAT T-3'. For strategy 2 (detection of *leader*RNA+cRNA, cRNA and Renilla mRNA+cRNA in total RNA from MG-transfected cells), the same RT primers as for strategy 1 were used to detect *leader*RNA+cRNA and cRNA only. Renilla mRNA+cRNA were reverse-transcribed with the RT primer luc (-): 5'-AGA ACC ATT ACC AGA TTT GCC TGA-3' hybridizing to the coding region of the Renilla luciferase reporter gene. Additionally, for normalization, Firefly luciferase mRNA (encoded by the cotransfected pGL4.13 plasmid) was reverse-transcribed using a Random Hexamer Primer set (TFS). RT-PCR strategy 3 (detection of vRNA, cRNA and Renilla mRNA in total RNA from MG 3'-end mutants) was performed as described previously (Bach et al., 2019a), utilizing the RT primer luc (+) 5'-GGC CTC TTC TTA TTT ATG GCG A-3' for the detection of negative-strand vRNA, primer luc (-) [see above] for RT of positive-stranded mRNA+cRNA, and primer RT\_cRNA s3 5'-CAG TCC TGC CTT TTC TTT TAA TTT TAT C-3' for specific detection of the cRNA trailer region.

### *Quantitative real-time (qRT-) PCR of cDNA derived from EBOV-infected cells (strategy 1)*

~ 40 ng of the prepared cDNA (4 µL of a 1:5 dilution) were used for the qRT-PCR reaction that was performed in technical duplicates on a BioRadiQ™5 system (BioRad) using the 2x ABsolute QPCR Capillary Mix, SYBR Green (TFS) and the respective forward and reverse primers (5 µM each) in a total volume of 20 µL. For amplification of

*leader*RNA+cRNA, primers RT\_leader+cRNA (see above) and qPCR\_leader+cRNA: 5'-GCC GGA CAC ACA AAA AGA AAG AA-3' were combined, yielding a 66 bp product. cRNA was specifically amplified using primers RT\_cRNA s1/s2 (see above) and qPCR\_leader+cRNA (see above), resulting in a product of 139 bp. NP mRNA was amplified using primers RT\_NP mRNA (see above) and qPCR\_NP mRNA 5'-GTA ACT CAA TAT TCT AAC TAG CGA TTT ATC TAA ATT AAA TTA CA -3' (product length 141 bp). The temperature profile was chosen as follows: 15 min 95 °C; 30 cycles of 95 °C 10 s / 60 °C 20 s / 72 °C 15 s. Calculation of the mRNA:*leader*RNA ratio was performed as depicted in Fig. S3.

#### *qRT-PCR of cDNA derived from MG-transfected cells (strategies 2 and 3)*

cDNAs were diluted 1:10 and 2 µL of the dilution (~ 5 ng) were used in the respective PCR reaction; qRT\_PCR was performed in technical duplicates in a total volume of 10 µL on a QuantStudio3™ Real-Time PCR System (TFS) using the PowerUp™ SYBR™ Green Master Mix 2x (TSF). PCR conditions were chosen according to the manufacturer's fast cycling mode protocol in 40 repeating cycles. RNA levels were quantified using the  $2^{-\Delta\Delta C_T}$  algorithm analogous to (Bach et al., 2019a). Primers were used as follows: in both strategies (2 and 3; Fig. 3B; Fig. S2) amplification of Firefly luciferase mRNA was performed using the validated (Bach et al., 2019a) primers qPCR\_FF\_fwd: 5'-CGT GCA AAA GAA GCT ACC G-3' and qPCR\_FF\_rev: 5'-GGT GGC AAA TGG GAA GTC AC-3' (108 bp product). Primers RT\_leader+cRNA and qPCR\_leader+cRNA (see above) were used in strategy 2 for both, amplification of *leader*RNA+cRNA and cRNA alone (66 bp product). Specificity for cRNA was achieved by usage of the primer in the RT step that only reverse-transcribed antisense (c)RNAs longer than ~ 120 nt and thus exceeding the maximal length of *leader*RNAs (~ 85 nt). For amplification of Renilla mRNA+cRNA, primers luc (-) and luc (+) were used (see above), yielding a product of 112 bp. Primer efficiencies were calculated according to Ramakers et al. (2003) for the wt NP MG based on 3 independent experiments with 2 to 3 replicates, resulting in  $2.00 \pm 0.01$  (standard deviation) for cRNA alone,  $2.00 \pm 0.02$  for *leader*RNA+cRNA and  $2.00 \pm 0.01$  for mRNA+cRNA. In strategy 3 (Bach et al., 2019a), vRNA and mRNA+cRNA were amplified using primers luc (+) and luc (-) (see above). Specific amplification of cRNA was obtained with primers RT\_cRNA s3 (see above) and qPCR\_cRNA s3: 5'-CGG TGA TAG CCT TAA TCT TTG TG-3' (118 bp product).

#### *Quantification and statistical analysis*

Statistical analysis was performed using GraphPad Prism version 8.1.1 for Windows. For details (such as *p* values), see figure legends. A statistical significance level of 0.05 was

chosen for analysis by the unpaired parametric Welch's  $t$  test or the non-parametric Mann Whitney test.

## **FUNDING**

This work was supported by the German Research Foundation (DFG), grant CRC 1021, to S. Becker and R. K. Hartmann.

## **ACKNOWLEDGEMENTS**

We like to acknowledge technical support by Astrid Herwig (cell culture) and Arnold Grünweller for fruitful discussions.

## REFERENCES

- Albariño, C. G., Guerrero, L. W., Chakrabarti, A. K., and Nichol, S. T. (2018). Transcriptional analysis of viral mRNAs reveals common transcription patterns in cells infected by five different filoviruses. *PLoS One* 13, 1–13. doi:10.1371/journal.pone.0201827.
- Biedenkopf, N., Hartlieb, B., Hoenen, T., and Becker, S. (2013). Phosphorylation of Ebola Virus VP30 Influences the Composition of the Viral Nucleocapsid Complex. *J. Biol. Chem.* 288, 11165–11174. doi:10.1074/jbc.m113.461285.
- Biedenkopf, N., and Hoenen, T. (2017). “Modeling the Ebolavirus Life Cycle with Transcription and Replication-Competent Viruslike Particle Assays,” in *Methods in molecular biology* (Clifton, N.J.), 119–131. doi:10.1007/978-1-4939-7116-9\_9.
- Biedenkopf, N., Lier, C., and Becker, S. (2016a). Dynamic Phosphorylation of VP30 Is Essential for Ebola Virus Life Cycle. *J. Virol.* 90, 4914–4925. doi:10.1128/jvi.03257-15.
- Biedenkopf, N., Schlereth, J., Grünweller, A., Becker, S., and Hartmann, R. K. (2016b). RNA Binding of Ebola Virus VP30 Is Essential for Activating Viral Transcription. *J. Virol.* 90, 7481–7496. doi:10.1128/jvi.00271-16.
- Bitko, V., Musiyenko, A., Bayfield, M. A., Maraia, R. J., and Barik, S. (2008). Cellular La protein shields nonsegmented negative-strand RNA viral leader RNA from RIG-I and enhances virus growth by diverse mechanisms. *J. Virol.* 82, 7977–87. doi:10.1128/JVI.02762-07.
- Blakqori, G., Kochs, G., Haller, O., and Weber, F. (2003). Functional L polymerase of La Crosse virus allows in vivo reconstitution of recombinant nucleocapsids. *J. Gen. Virol.* 84, 1207–1214. doi:10.1099/vir.0.18876-0.
- Blumberg, B. M., Leppert, M., and Kolakofsky, D. (1981). Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication. *Cell* 23, 837–45. doi:10.1016/0092-8674(81)90448-7.
- Brauburger, K., Boehmann, Y., Krähling, V., and Mühlberger, E. (2016). Structure and Regulatory Elements on Gene Expression and Polymerase Scanning Behavior. *J. Virol.* 90, 1898–1909. doi:10.1128/JVI.02341-15.Editor.
- Calain, P., Monroe, M. C., and Nichol, S. T. (1999). Ebola Virus Defective Interfering Particles and Persistent Infection. *Virology* 262, 114–128. doi:10.1006/VIRO.1999.9915.
- Calain, P., and Roux, L. (1993). The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *J. Virol.* 67, 4822–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8392616> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC237869>.
- Chuang, J. L., and Perrault, J. (1997). Initiation of vesicular stomatitis virus mutant polR1 transcription internally at the N gene in vitro. *J. Virol.* 71, 1466–75. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8995672> [Accessed September 7, 2019].
- Cordey, S., and Roux, L. (2006). Transcribing paramyxovirus RNA polymerase engages the template at its 3' extremity. *J. Gen. Virol.* 87, 665–672. doi:10.1099/vir.0.81353-

0.

- Cordey, S., and Roux, L. (2007). Further characterization of a paramyxovirus transcription initiation signal: Search for required nucleotides upstream and importance of the N phase context. *J. Gen. Virol.* 88, 1555–1564. doi:10.1099/vir.0.82701-0.
- Cowton, V. M., McGivern, D. R., and Fearn, R. (2006). Unravelling the complexities of respiratory syncytial virus RNA synthesis. *J. Gen. Virol.* 87, 1805–1821. doi:10.1099/vir.0.81786-0.
- Cox, R., Pickar, A., Qiu, S., Tsao, J., Rodenburg, C., Dokland, T., et al. (2014). Structural studies on the authentic mumps virus nucleocapsid showing uncoiling by the phosphoprotein. *Proc. Natl. Acad. Sci.* 111, 15208–15213. doi:10.1073/pnas.1413268111.
- Deflubé, L. R., Cressey, T. N., Hume, A. J., Olejnik, J., Haddock, E., and Feldmann, F. (2019). Ebolavirus polymerase uses an unconventional genome replication mechanism. 116, 8535–8543. doi:10.1073/pnas.1815745116.
- Fearn, R., and Plemper, R. K. (2017). Polymerases of paramyxoviruses and pneumoviruses. *Virus Res.* 234, 87–102. doi:10.1016/j.virusres.2017.01.008.
- Garcin, D., and Kolakofsky, D. (1992). Tacaribe arenavirus RNA synthesis in vitro is primer dependent and suggests an unusual model for the initiation of genome replication. *J. Virol.* 66, 1370–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1738196> [Accessed September 7, 2019].
- Ghosh, A., Nayak, R., and Shaila, M. S. (1996). Synthesis of leader RNA and editing of P mRNA during transcription by rinderpest virus. *Virus Res.* 41, 69–76. doi:10.1016/0168-1702(95)01276-1.
- Hanley, L. L., McGivern, D. R., Teng, M. N., Djang, R., Collins, P. L., and Fearn, R. (2010). Roles of the respiratory syncytial virus trailer region: effects of mutations on genome production and stress granule formation. *Virology* 406, 241–52. doi:10.1016/j.virol.2010.07.006.
- Hinzman, E. E., Barr, J. N., and Wertz, G. W. (2002). Identification of an upstream sequence element required for vesicular stomatitis virus mRNA transcription. *J. Virol.* 76, 7632–41. doi:10.1128/jvi.76.15.7632-7641.2002.
- Hoch, P. G., Schlereth, J., Lechner, M., and Hartmann, R. K. (2016). *Bacillus subtilis* 6S-2 RNA serves as a template for short transcripts in vivo. *RNA* 22, 614–622. doi:10.1261/rna.055616.115.
- Hoenen, T., Jung, S., Herwig, A., Groseth, A., and Becker, S. (2010). Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription. *Virology* 403, 56–66. doi:10.1016/j.virol.2010.04.002.
- Hoffman, M. A., and Banerjee, A. K. (2000). Precise Mapping of the Replication and Transcription Promoters of Human Parainfluenza Virus Type 3. *Virology* 269, 201–211. doi:10.1006/VIRO.2000.0223.
- Hoffmann, S., Otto, C., Kurtz, S., Sharma, C. M., Khaitovich, P., Vogel, J., et al. (2009). Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput. Biol.* 5, e1000502.

doi:10.1371/journal.pcbi.1000502.

- Honda, T., Sofuku, K., Kojima, S., Yamamoto, Y., Ohtaki, N., and Tomonaga, K. (2017). Linkage between the leader sequence and leader RNA production in Borna disease virus-infected cells. *Virology*. doi:10.1016/j.virol.2017.07.011.
- Horikami, S. M., and Moyer, S. A. (1991). Synthesis of leader RNA and editing of the P mRNA during transcription by purified measles virus. *J. Virol.* 65, 5342–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1895388> [Accessed September 4, 2019].
- Hume, A. J., and Mühlberger, E. (2019). Distinct Genome Replication and Transcription Strategies within the Growing Filovirus Family. *J. Mol. Biol.* doi:10.1016/j.jmb.2019.06.029.
- Iseni, F., Garcin, D., Nishio, M., Kedersha, N., Anderson, P., and Kolakofsky, D. (2002). Sendai virus trailer RNA binds TIAR, a cellular protein involved in virus-induced apoptosis. *EMBO J.* 21, 5141. doi:10.1093/EMBOJ/CDF513.
- Kolakofsky, D. (2016). Paramyxovirus RNA synthesis , mRNA editing , and genome hexamer phase : A review. *Virology* 498, 94–98. doi:10.1016/j.virol.2016.08.018.
- Kruse, T., Biedenkopf, N., Hertz, E. P. T., Dietzel, E., Stalman, G., López-Méndez, B., et al. (2018). The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. *Mol. Cell* 69, 136–145.e6. doi:10.1016/j.molcel.2017.11.034.
- Kurilla, M. G., Stone, H. O., and Keene, J. D. (1985). RNA sequence and transcriptional properties of the 3' end of the Newcastle disease virus genome. *Virology* 145, 203–212. doi:10.1016/0042-6822(85)90154-0.
- Le Mercier, P., and Kolakofsky, D. (2019). Bipartite promoters and RNA editing of paramyxoviruses and filoviruses. *RNA* 25, 279–285. doi:10.1261/rna.068825.118.
- Leppert, M., Rittenhouse, L., Perrault, J., Summers, D. F., and Kolakofsky, D. (1979). Plus and minus strand leader RNAs in negative strand virus-infected cells. *Cell* 18, 735–747. doi:10.1016/0092-8674(79)90127-2.
- Li, T., and Pattnaik, A. K. (1999). Overlapping signals for transcription and replication at the 3' terminus of the vesicular stomatitis virus genome. *J. Virol.* 73, 444–52. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9847350> [Accessed August 29, 2019].
- Louber, J., Kowalinski, E., Bloyet, L. M., Brunel, J., Cusack, S., and Gerlier, D. (2014). RIG-I self-oligomerization is either dispensable or very transient for signal transduction. *PLoS One* 9, e108770. doi:10.1371/journal.pone.0108770.
- Malvy, D., McElroy, A. K., de Clerck, H., Günther, S., and van Griensven, J. (2019). Ebola virus disease. *Lancet* 393, 936–948. doi:10.1016/S0140-6736(18)33132-5.
- Marq, J.-B., Kolakofsky, D., and Garcin, D. (2010). Unpaired 5' ppp-nucleotides, as found in arenavirus double-stranded RNA panhandles, are not recognized by RIG-I. *J. Biol. Chem.* 285, 18208–16. doi:10.1074/jbc.M109.089425.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17, 10. doi:10.14806/ej.17.1.200.
- Mehedi, M., Falzarano, D., Seebach, J., Hu, X., Carpenter, M. S., Schnittler, H.-J., et al.

- (2011). A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J. Virol.* 85, 5406–14. doi:10.1128/JVI.02190-10.
- Messaoudi, I., Amarasinghe, G. K., and Basler, C. F. (2015). Filovirus pathogenesis and immune evasion: Insights from Ebola virus and Marburg virus. *Nat. Rev. Microbiol.* 13, 663–676. doi:10.1038/nrmicro3524.
- Mühlberger, E. (2007). Filovirus replication and transcription. *Future Virol.* 2, 205–215. doi:10.2217/17460794.2.2.205.
- Mühlberger, E., Lötfering, B., Klenk, H. D., and Becker, S. (1998). Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. *J. Virol.* 72, 8756–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9765419> [Accessed August 27, 2019].
- Mühlberger, E., Weik, M., Volchkov, V. E., Klenk, H. D., and Becker, S. (1999). Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J. Virol.* 73, 2333–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9971816> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC104478>.
- Murphy, S. K., and Parks, G. D. (1999). RNA replication for the paramyxovirus simian virus 5 requires an internal repeated (CGNNNN) sequence motif. *J. Virol.* 73, 805–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9847393> [Accessed August 30, 2019].
- Noton, S. L., and Fearn, R. (2015). Initiation and regulation of paramyxovirus transcription and replication. *Virology* 479–480, 545–554. doi:10.1016/j.virol.2015.01.014.
- Ogino, T., and Green, T. J. (2019). RNA Synthesis and Capping by Non-segmented Negative Strand RNA Viral Polymerases: Lessons From a Prototypic Virus. *Front. Microbiol.* 10, 1490. doi:10.3389/fmicb.2019.01490.
- Oh, S., Onomoto, K., Wakimoto, M., and Onoguchi, K. (2016). Leader-Containing Uncapped Viral Transcript Activates RIG-I in Antiviral Stress Granules. 1–22. doi:10.1371/journal.ppat.1005444.
- Olsen, M. E., Cressey, T. N., Mühlberger, E., and Connor, J. H. (2018). Differential Mechanisms for the Involvement of Polyamines and Hypusinated eIF5A in Ebola Virus Gene Expression. *J. Virol.* 92, e01260-18. doi:10.1128/JVI.01260-18.
- Patton, J. T., Davis, N. L., and Wertz, G. W. (1984). N protein alone satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. *J. Virol.* 49, 303–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6319730> [Accessed September 4, 2019].
- Perez, J. T., Varble, A., Sachidanandam, R., Zlatev, I., Manoharan, M., García-Sastre, A., et al. (2010). Influenza A virus-generated small RNAs regulate the switch from transcription to replication. *Proc. Natl. Acad. Sci. U. S. A.* 107, 11525–30. doi:10.1073/pnas.1001984107.
- Perrault, J., Clinton, G. M., and McClure, M. A. (1983). RNP template of vesicular stomatitis virus regulates transcription and replication functions. *Cell* 35, 175–185. doi:10.1016/0092-8674(83)90220-9.

- Plumet, S., Herschke, F., Bourhis, J. M., Valentin, H., Longhi, S., and Gerlier, D. (2007). Cytosolic 5'-triphosphate ended viral leader transcript of measles virus as activator of the RIG I-mediated interferon response. *PLoS One* 2, 1–11. doi:10.1371/journal.pone.0000279.
- Ramakers, C., Ruijter, J. M., Lekanne Deprez, R. H., and Moorman, A. F. M. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* 339, 62–66. doi:10.1016/S0304-3940(02)01423-4.
- Rougeron, V., Feldmann, H., Grard, G., Becker, S., and Leroy, E. M. (2015). Ebola and Marburg haemorrhagic fever. *J. Clin. Virol.* 64, 111–119. doi:10.1016/j.jcv.2015.01.014.
- Sanchez, A., Kiley, M. P., Holloway, B. P., and Auperin, D. D. (1993). Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res.* 29, 215–240. doi:10.1016/0168-1702(93)90063-S.
- Schlee, M., Roth, A., Hornung, V., Hagmann, C. A., Wimmenauer, V., Barchet, W., et al. (2009). Recognition of 5' Triphosphate by RIG-I Helicase Requires Short Blunt Double-Stranded RNA as Contained in Panhandle of Negative-Strand Virus. *Immunity* 31, 25–34. doi:10.1016/j.immuni.2009.05.008.
- Schlereth, J., Grünweller, A., Biedenkopf, N., Becker, S., and Hartmann, R. K. (2016). RNA binding specificity of Ebola virus transcription factor VP30. *RNA Biol.* 13, 783–798. doi:10.1080/15476286.2016.1194160.
- Shabman, R. S., Jabado, O. J., Mire, C. E., Stockwell, T. B., Edwards, M., Mahajan, M., et al. (2014). Deep Sequencing Identifies Noncanonical Editing of Ebola and Marburg Virus RNAs in Infected Cells. 5, 1–11. doi:10.1128/mBio.02011-14. Editor.
- Sugita, Y., Matsunami, H., Kawaoka, Y., Noda, T., and Wolf, M. (2018). Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex at 3.6 Å resolution. *Nature*. doi:10.1038/s41586-018-0630-0.
- Sztuba-Solinska, J., Diaz, L., Kumar, M. R., Kolb, G., Wiley, M. R., Jozwick, L., et al. (2016). A small stem-loop structure of the Ebola virus trailer is essential for replication and interacts with heat-shock protein A8. *Nucleic Acids Res.* 44, 9831–9846. doi:10.1093/nar/gkw825.
- Tapparel, C., Maurice, D., and Roux, L. (1998). The activity of Sendai virus genomic and antigenomic promoters requires a second element past the leader template regions: a motif (GNNNNN)3 is essential for replication. *J. Virol.* 72, 3117–28. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9525637> [Accessed August 28, 2019].
- The R Development Core Team (2011). R: A Language and Environment for Statistical Computing Reference Index. R Foundation for Statistical Computing Available at: <https://stat.ethz.ch/pipermail/r-help/2008-May/161481.html> [Accessed November 29, 2019].
- Tremaglio, C. Z., Noton, S. L., Deflube, L. R., and Fearn, R. (2013). Respiratory Syncytial Virus Polymerase Can Initiate Transcription from Position 3 of the Leader Promoter. *J. Virol.* 87, 3196–3207. doi:10.1128/JVI.02862-12.
- Umbach, J. L., Yen, H.-L., Poon, L. L. M., and Cullen, B. R. (2010). Influenza A virus expresses high levels of an unusual class of small viral leader RNAs in infected



- cells. MBio 1. doi:10.1128/mBio.00204-10.
- Vidal, S., and Kolakofsky, D. (1989). Modified model for the switch from Sendai virus transcription to replication. *J. Virol.* 63, 1951–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2539496> [Accessed August 30, 2019].
- Volchkov, V. E., Volchkova, V. A., Chepurinov, A. A., Blinov, V. M., Dolnik, O., Netesov, S. V, et al. (1999). Characterization of the L gene and 5' trailer region of Ebola virus. *J. Gen. Virol.* 80, 355–362. doi:10.1099/0022-1317-80-2-355.
- Vulli  moz, D., Cordey, S., Mottet-Osman, G., and Roux, L. (2005). Nature of a paramyxovirus replication promoter influences a nearby transcription signal. *J. Gen. Virol.* 86, 171–180. doi:10.1099/vir.0.80435-0.
- Wan, W., Kolesnikova, L., Clarke, M., Koehler, A., Noda, T., Becker, S., et al. (2017). Structure and assembly of the Ebola virus nucleocapsid. *Nature* 551, 394–397. doi:10.1038/nature24490.
- Weik, M., Enterlein, S., Schlenz, K., and M  hlberger, E. (2005). The Ebola Virus Genomic Replication Promoter Is Bipartite and Follows the Rule of Six. *J. Virol.* 79, 10660–10671. doi:10.1128/jvi.79.16.10660-10671.2005.
- Weik, M., Modrof, J., Klenk, H.-D., Becker, S., and M  hlberger, E. (2002). Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J. Virol.* 76, 8532–9. doi:10.1128/JVI.76.17.8532.
- Whelan, S. P. J., Barr, J. N., and Wertz, G. W. (2004). Transcription and replication of nonsegmented negative-strand RNA viruses. *Curr. Top. Microbiol. Immunol.* 283, 61–119. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15298168>.
- Whelan, S. P., and Wertz, G. W. (1999). Regulation of RNA synthesis by the genomic termini of vesicular stomatitis virus: identification of distinct sequences essential for transcription but not replication. *J. Virol.* 73, 297–306. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9847333> [Accessed August 3, 2019].
- Xu, W., Luthra, P., Wu, C., Batra, J., Leung, D. W., Basler, C. F., et al. (2017). Ebola virus VP30 and nucleoprotein interactions modulate viral RNA synthesis. *Nat. Commun.* 8, 1–11. doi:10.1038/ncomms15576.

## FIGURE LEGENDS

**Fig. 1: (A)** EBOV genome with 3'-leader and 5'-trailer regions relevant to viral replication and encoding 7 structural mRNAs whose abundance decreases gradually from the first NP mRNA to the last L mRNA. Secondary structure formation potential of the 7 transcription start signals (TSS, in cyan) is depicted above the schematic genome. Potential 5'-UTR structures antisense to the genomic TSS-secondary structures are not shown. White boxes indicate the open reading frames for EBOV proteins; 5'- and 3'-UTRs are marked by light gray boxes. Dark gray areas depict the position of the predicted secondary structures on genomic RNA level. Locations of leader and trailer promoters are indicated by small letter p. mRNAs of the 7 Ebola genes are depicted as horizontal lines with round circles at their respective 5' ends, marking 5' cap structures. **(B)** Zoom into the genomic 3'-terminal (upper panel) and corresponding antigenomic (lower panel) sequence and secondary structure potential, including the 55 nt 3'-leader, replication promoter elements 1 and 2 (PE1, PE2, in green) and the NP TSS (cyan; a blue arrow marks the mRNA initiation site) which forms a hairpin structure (NP hairpin) together with a spacer region (orange; Schlereth et al., 2016; Weik et al., 2002, 2005). NP hairpin structure formation was suggested to regulate VP30-dependent transcription and to influence the VP30-NP interaction (Weik et al., 2002; Xu et al., 2017). Genomic RNA nucleotides are numbered with negative values starting at the 3'-terminal nt (position -1), vice versa, complementary antigenomic nucleotides are numbered using positive values. UN<sub>5</sub> hexamers (marked by bold pink 3'-U residues) either between genome nucleotides (nt) -51 and -80 or in PE2 (-81 to -128) were previously shown to be relevant to replication and transcription (Weik et al., 2005; Bach et al., 2019a); The 3'-terminal G of the genome vRNA is depicted as italic letter considering the recent finding that EBOV RNA polymerase initiates antigenome synthesis precisely opposite genome position -2 (red arrow), even with genome 3' end extensions of 3 nt (Deflubé et al., 2019). Formation of a genome 3'-terminal secondary structure [leader (Le) hairpin] was hypothesized to be relevant in a back-priming mechanism for the non-templated addition of the 1<sup>st</sup> genome nt (Deflubé et al., 2019). Red bold letters in the vRNA sequence indicate nt positions whose sequence preservations have been shown to be important to replication (Weik et al., 2005). Genome nucleotides -54 and -55 (italic bold blue letters) were shown to be obligatory to both replication and transcription (Weik et al., 2005; Bach et al., 2019a).

**Fig. 2:** Transcriptome analysis of small RNAs (< 200 nt; A) or poly(A) RNAs (B) derived from EBOV-infected HuH7 cells using Next Generation Sequencing (NGS). **(A)** Analysis of transcript 3' ends (dark blue bars) in the small RNA library. Relative reads (y-axis)

+/- standard error of the mean (SEM) for each transcript length (x-axis; [nt]) are indicated as well as relative reads for certain length regions marked by dashed vertical lines. They derive from 3 biological replicates. Below the diagram, a graphical sketch of the viral antigenome is indicated to correlate transcript 3' ends to the respective antigenome region. **(B)** Relative read numbers (y-axis) +/- SEM of NP mRNA start sites derive from 2 independent experiments. x-axis: nucleotide (nt) position corresponding to the respective MG template RNA [sketch below bar diagram, including the major transcription initiation site at genome nt -56 (blue arrow)].

**Fig. 3:** Northern Blot (A) and qRT-PCR analysis (B-E) of total RNA derived from EBOV-infected HuH7 cells. **(A)** Northern Blot detection of *leader*RNA using a fully complementary digoxigenin-labeled *in-vitro*-transcribed RNA probe (nt 80-1 of the genomic RNA; schematically illustrated on the left; For further information see Material and Methods and Supplementary Table S1). A T7-transcribed reference RNA 2-78 served as positive control and size marker, T7-transcripts 56-158 and 154-1 (corresponding to the equivalent antigenome or genome sequence) were included as specificity controls. Total RNA from Mock (non-infected) cells was used as negative (neg.) control. Leader transcripts of ~ 60 -70 nts were detected 15 min and 30 min post exposure. **(B)** Schematic qRT-PCR primer setup 1 (s1) used for detection of *leader*RNA+cRNA (blue arrows), cRNA (red arrows), and NP mRNA (grey arrows); 5'-ends of the RT primers (dashed vertical lines and nt position) are indicated. **(C)** Exemplary qPCR amplification plot obtained by the three different primer sets for *leader*RNA+cRNA, cRNA or NP mRNA. Y-axis: relative fluorescence signal (RFU) per PCR cycle; mock = non-infected cells. **(D)** Quantification of the NP mRNA : *leader*RNA ratio for 3 biological (biol.) replicates with 2 technical replicates each and the resulting mean ratio +/- SEM. An exemplary calculation for one of the replicates (highlighted in yellow) is shown in Supplementary Fig. S1.

**Fig. 4:** **(A)** Quantification of *leader(le)*RNA synthesis in MG-transfected HEK293 cells by 2-step qRT-PCR. The qRT-PCR setup (s2) is schematically depicted in Fig. S2. Relative levels ( $2^{-\Delta CT}$  values; normalized to Firefly luciferase mRNA) of mRNA+cRNA, *le*RNA+cRNA, and cRNA in the presence (+) or absence (-) of the viral polymerase L are shown (mean  $2^{-\Delta CT} = \sim 300$  for mRNA+cRNA; mean  $2^{-\Delta CT} = \sim 5$  for *leader*RNA+cRNA, and mean  $2^{-\Delta CT} = \sim 2$  for cRNA alone). Mean values +/- SEM derive from 5 independent experiments (exp.) with 2 (1 of 5 exp.) or 3 (4 of 5 exp.) technical replicates each. **(B)** Leader transcript starting nucleotides determined by NGS for i) EBOV-infected HuH7 cells, lane 1; ii) wt MG-transfected HEK293 cells, lane 2; iii) a mock (non-infected/non-transfected) HEK293 control spiked with synthetic 65-meric *le*RNA, lane 3 (sequence in

table S2). For detailed information on color code see legend next to bar diagram. The number of biological replicates (n) is indicated above the respective lane.

**Fig. 5:** Analysis of replication-competent (RC) and replication-deficient (RD) MGs with mutated 3' ends. **(A)** 3' sequence of the wt MG and 3 mutant MGs which either lack the first 3' nt ( $\Delta 1$ ), the first two ( $\Delta 2$ ) 3' nts or harbor an additional 3'-G nt (+G). **(B-C)** Corresponding reporter gene assays of MGs illustrated in (A). Activity values of mutant MGs (light blue bars) are shown as means  $\pm$  SEM of 3 independent experiments with 3 technical replicates each relative to the native 3'-leader MG (grey bars; wt = 100%). As a negative control, the plasmid encoding the L gene was omitted during transfection (-L; black bars). **(D-F)** Corresponding 2-step qRT-PCR of RC MG samples using the same cells as in panel B [qRT-PCR setup according to Bach et al. (2019a)]. Color code as in legend to panel B-C. Mean  $2^{-\Delta\Delta C_T}$  values  $\pm$  SEM of viral mRNA (D), cRNA (E) and vRNA (F) derive from 3 independent experiments with 3 technical replicates each. \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001; \*\*\*\*p < 0.0001 (unpaired Welch's *t* test).

**Fig. 6:** Analysis of alternative gene end (GE) signals in the genome 3' sequence using a bicistronic (bici) MG system reflecting the EBOV NP-VP35 gene border. **(A)** Bici MGs encode a Renilla (light blue box) and Firefly (light red box) luciferase. Color code and genome scheme are essentially identical to Fig. 1. IGR, native NP-VP35 intergenic region (orange). In derivatives (structure 2 - 4 from top) of the wt NP-NP MG the native NP GE sequence (GEmut., light green boxes) was mutated (red vertical lines, illustrated in the box at the bottom). The native VP35 5'-UTR hairpin at the 2<sup>nd</sup> cistron was exchanged with i) a 2<sup>nd</sup> NP 5'-UTR hairpin (wt NP-NP, NPGEmut.-NP), ii) the leader(*le*) + NP hairpin (NPGEmut.-*le*NP) or iii) the VP40 5'-UTR hairpin which harbors a 2<sup>nd</sup> intact GE signal (GE2) and serves as a positive control. **(B-C)** Corresponding Renilla (B) and Firefly (C) luciferase reporter gene assays were used for indirect determination of mRNAs transcribed from the 1<sup>st</sup> or 2<sup>nd</sup> cistron. Activity values of the wt NP-NP MG were set to 100 %. -L, negative control in which the plasmid encoding L was omitted during transfection. Mean values  $\pm$  SEM derive from 3 independent experiments with at least 2 technical replicates each. \*p < 0.05; \*\*\*\*p < 0.0001; n.s., not significant (unpaired Welch's *t* test).

**Fig. 7:** *Leader*RNA synthesis in the presence of absence of VP30. **(A-C)** 2-step strand-specific qRT-PCR (strategy s2; Supplementary Fig. S2) quantification of mRNA+cRNA (A), cRNA (B) or *leader(le)*RNA+cRNA (C) using the  $2^{-\Delta\Delta C_T}$  algorithm analogous to Bach et al., (2019a). HEK293 cells were transfected with the components

of the replication-competent (RC) MG assay, with or without pCAGGS VP30 or pCAGGS L, as indicated. Mean  $2^{-\Delta\Delta C_T}$  values  $\pm$  SEM derive from 5 independent experiments (exp.) with 3 (4 of 5 exp.) or 2 (1 of 5 exp.) technical replicates. \* $p < 0.05$ ; \*\*\*\* $p < 0.0001$  (unpaired Welch's  $t$  test). **(D)** Relative mean  $2^{-\Delta C_T}$  values  $\pm$  SEM of *le*RNA+cRNA versus cRNA qRT-PCR primer sets corresponding to samples used in panel B and C.  $2^{-\Delta C_T}$  calculation was performed as described in legend to Fig. 4A or Material and Methods. The color code is equivalent to panel (A-C). \* $p < 0.05$  (unpaired Welch's  $t$  test).

**Fig. 8:** RNA Sequencing (RNA-Seq) analysis of leader transcript synthesis in RC MG systems, with (+) and without (-) cotransfection of pCAGGS VP30, as indicated. **(A)** Illustration of the wt NP MG and mutant derivatives either comprising a destabilized (NheI NP; Weik et al., 2002) or eliminated ( $\Delta 5'$  spacer; Bach et al. 2019b) NP HP structure on genomic (upper panel) or mRNA (lower panel) level. Minimum free energies (MFE;  $\Delta G$ ) of the centroid secondary structures were predicted by RNAfold by using the default parameters (<http://rna.tbi.univie.ac.at/cgi-bin/NAWebSuite/RNAfold.cgi>). Mutated nucleotides are marked by underlined black letters. The color code is essentially as in Fig. 1B. **(B-C)** Relative read numbers of *leader*RNAs (B) and Renilla mRNAs (C) obtained by RNA-Seq of small RNAs (< 200 nt, panel B) or mRNAs (poly(A) RNA, panel C) derived from MG samples illustrated in panel A. For details on sample and library preparation, Illumina Sequencing, and sequencing analyses see Supplementary Information and Table S3. **(D-E)** Renilla mRNA transcription start sites (TSS) obtained for  $\Delta 5'$  spacer MG systems with (+) and without (-) VP30, as indicated. Relative read numbers (Y-axis)  $\pm$  SEM derive from 3 biological replicates. X-axis: nucleotide (nt) position corresponding to the MG template RNA.

**Fig. 9:** Comparative RNA-Seq analysis of *leader*RNA 3' termination sites in EBOV-infected HuH7 cells **(A)** or MG-transfected HEK293 cells (wt NP, panel **(B)**; NheI NP MG, panel **(C)**,  $\Delta 5'$  spacer MG, panel **(D)**). MG variants are illustrated in Fig. 8A. X-axis: nt position corresponding to the (mini)genome RNA (illustrated below each bar diagram; color code similar to Fig. 8A). Relative read numbers  $\pm$  SEM (y-axis) derive from 3 (EBOV,  $\Delta 5'$  spacer MG) or 4 (wt MG, NheI NP MG) biological replicates. Dashed vertical lines mark regions < 60 nt, 60 to 69 nt; 70 to 79 nt, 80 to 89 nt and > 89 nt. Relative read numbers accounting for each region are indicated.

**Fig. 10:** qRT-PCR analysis of *leader*RNA synthesis from MG templates that do not obey the “rule of 6”. **(A)** Illustration of the wt NP RC MG and two mutant variants without spacer lengths of a multiple of 6 (29 nt). For detailed information on color code see

legend to Fig. 1B.  $\Delta$ , single nt deletion, as indicated. (**B-D**) qRT-PCR analysis (performed as described in Fig. 7 and Fig. S2) of MG variants described in panel A determining mRNA+cRNA (B), *le*RNA+cRNA (C) or cRNA (D) levels. Mean  $2^{-\Delta\Delta C_T}$  values derive from 3 independent experiments with 3 technical replicates each.

FIGURES

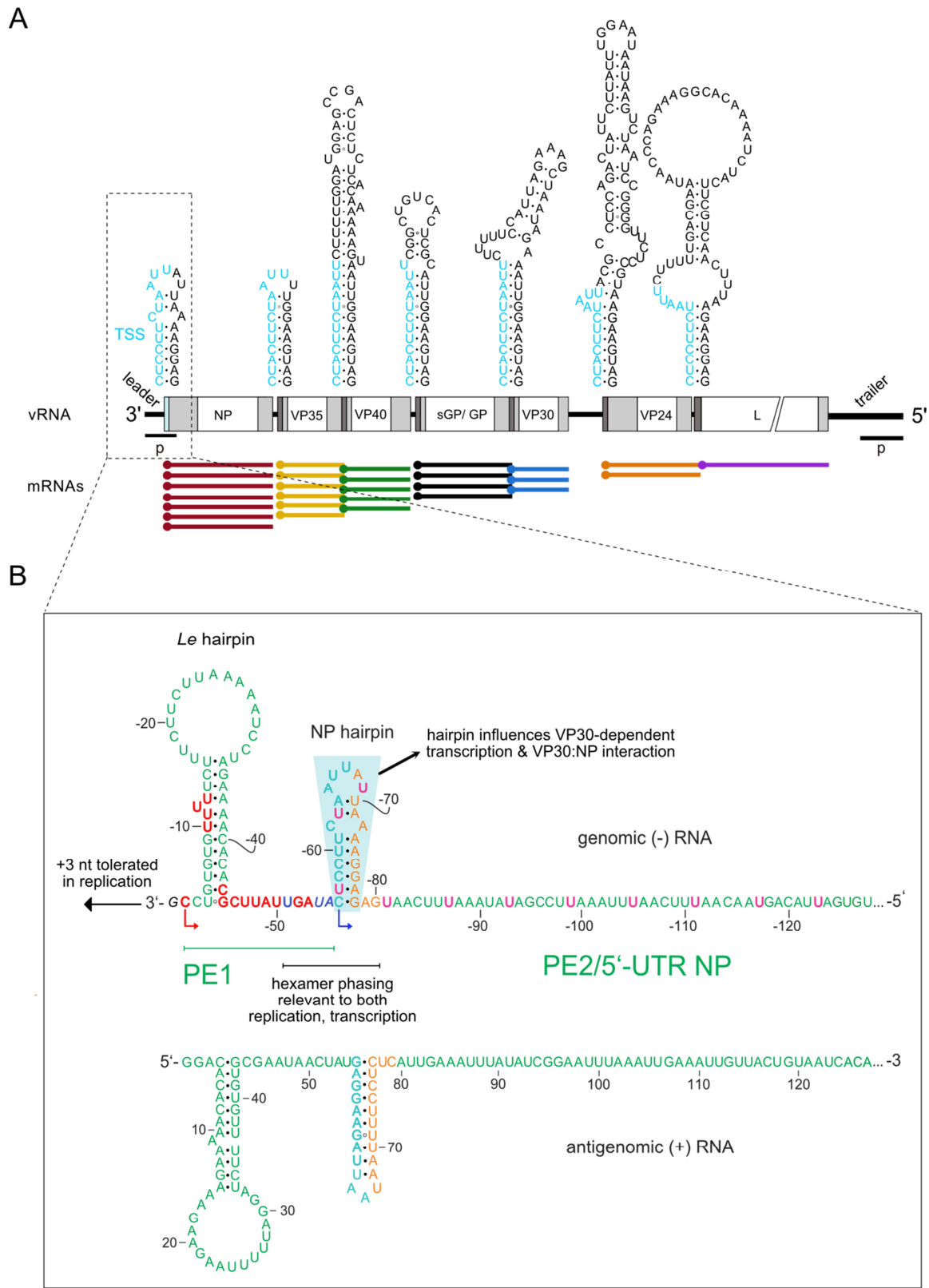


Fig.1

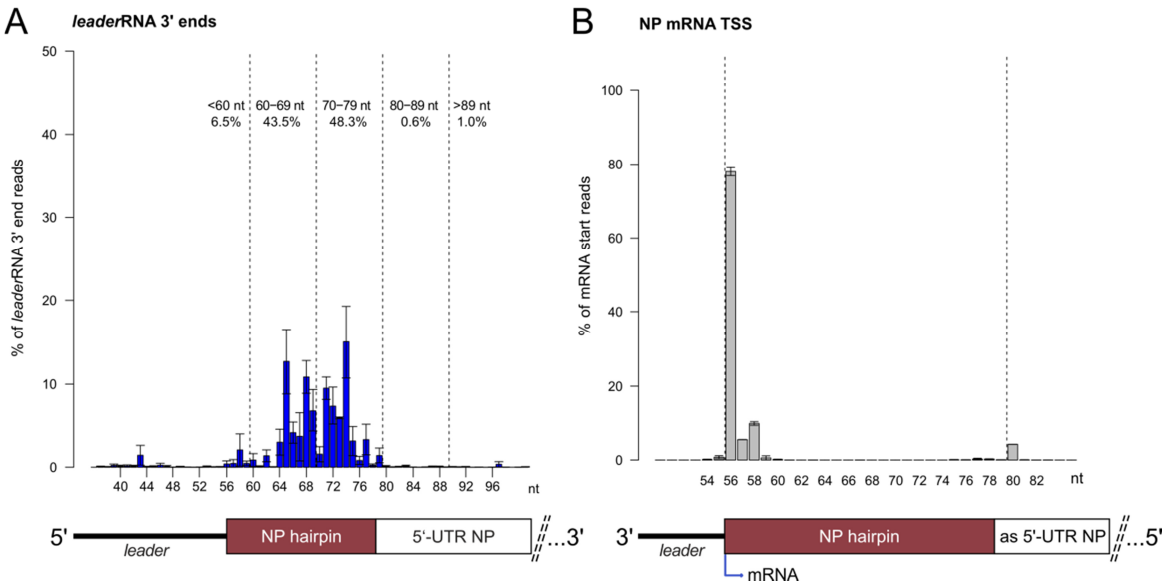


Fig. 2



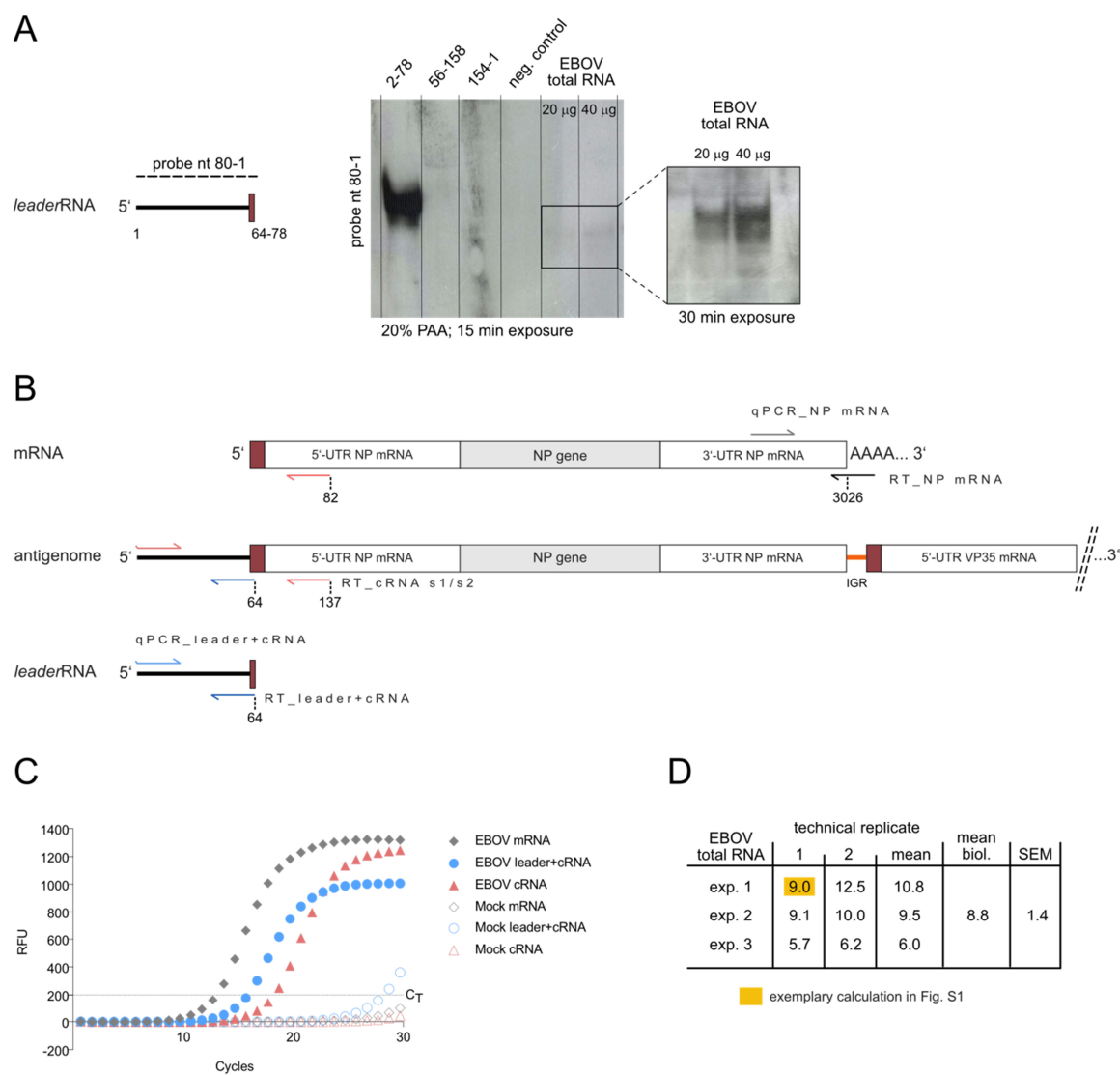


Fig. 3

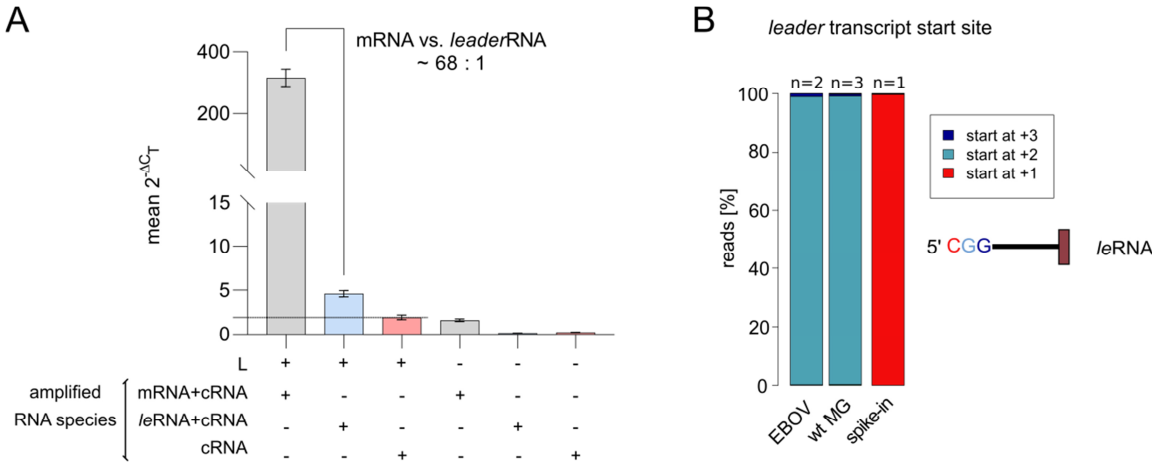


Fig. 4

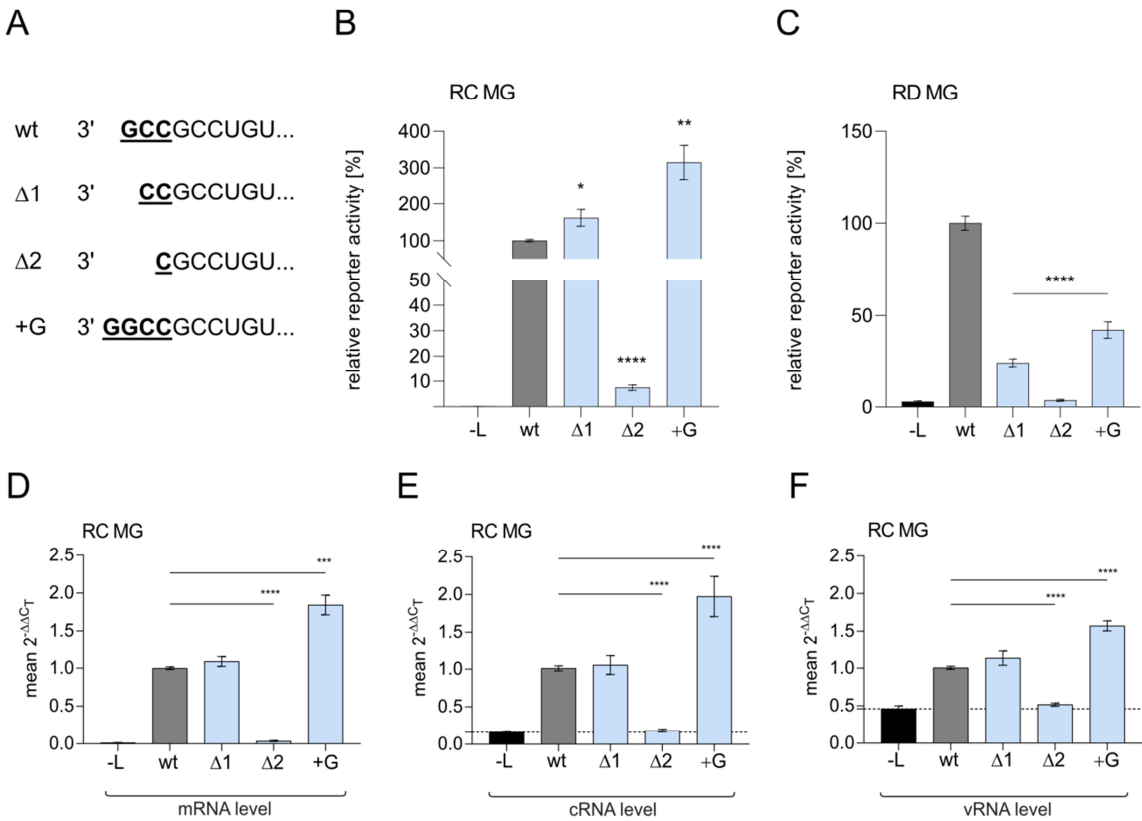


Fig. 5

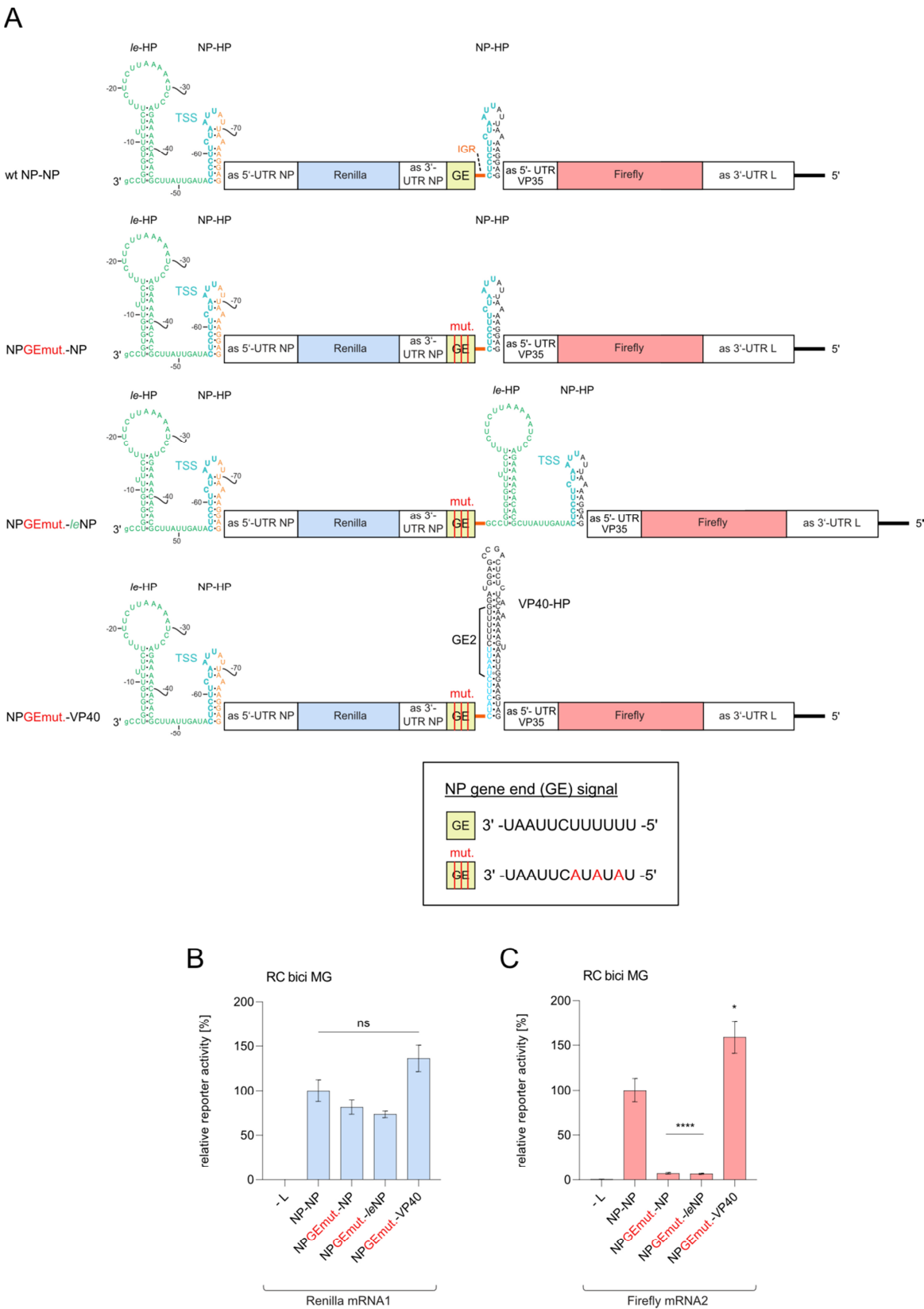


Fig. 6

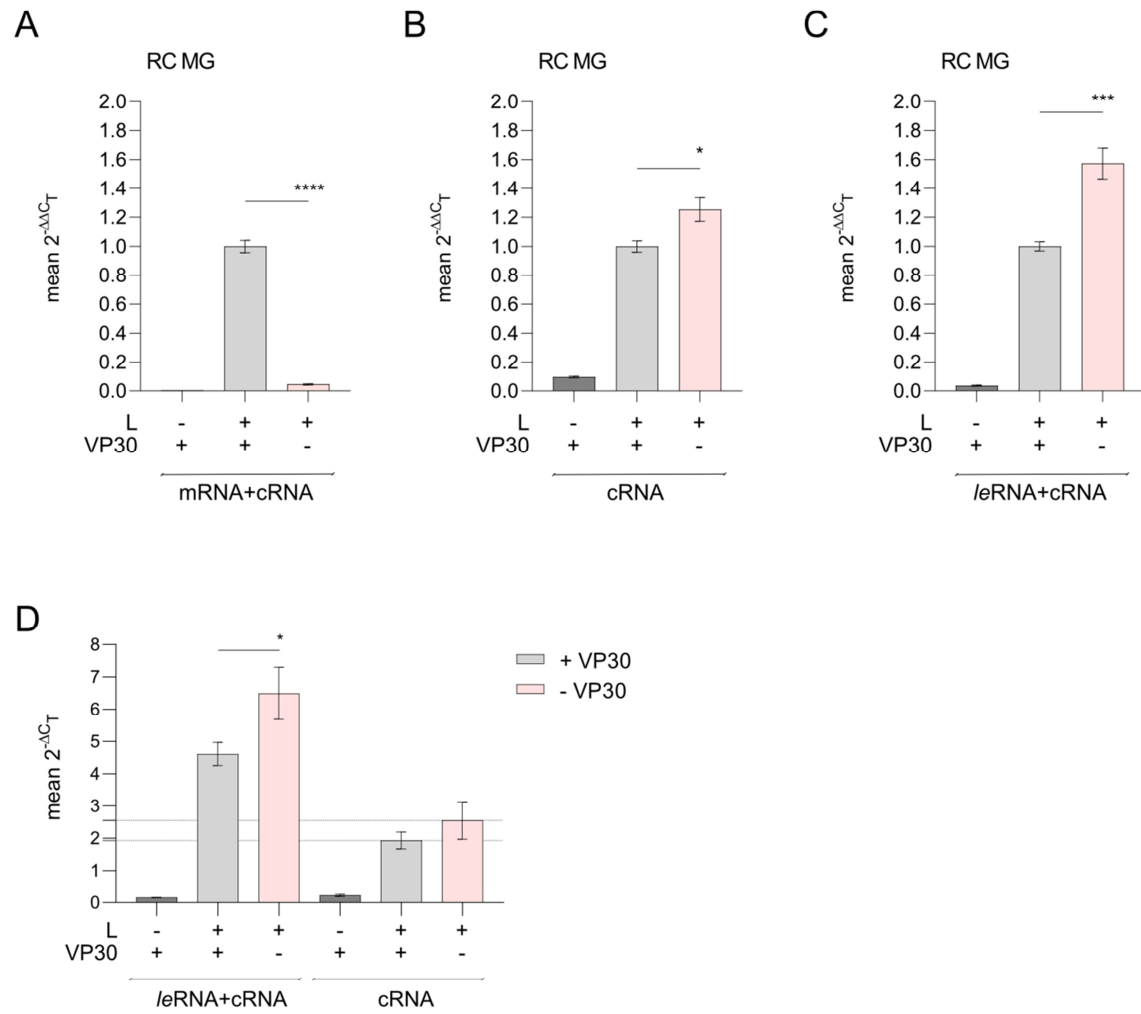


Fig. 7

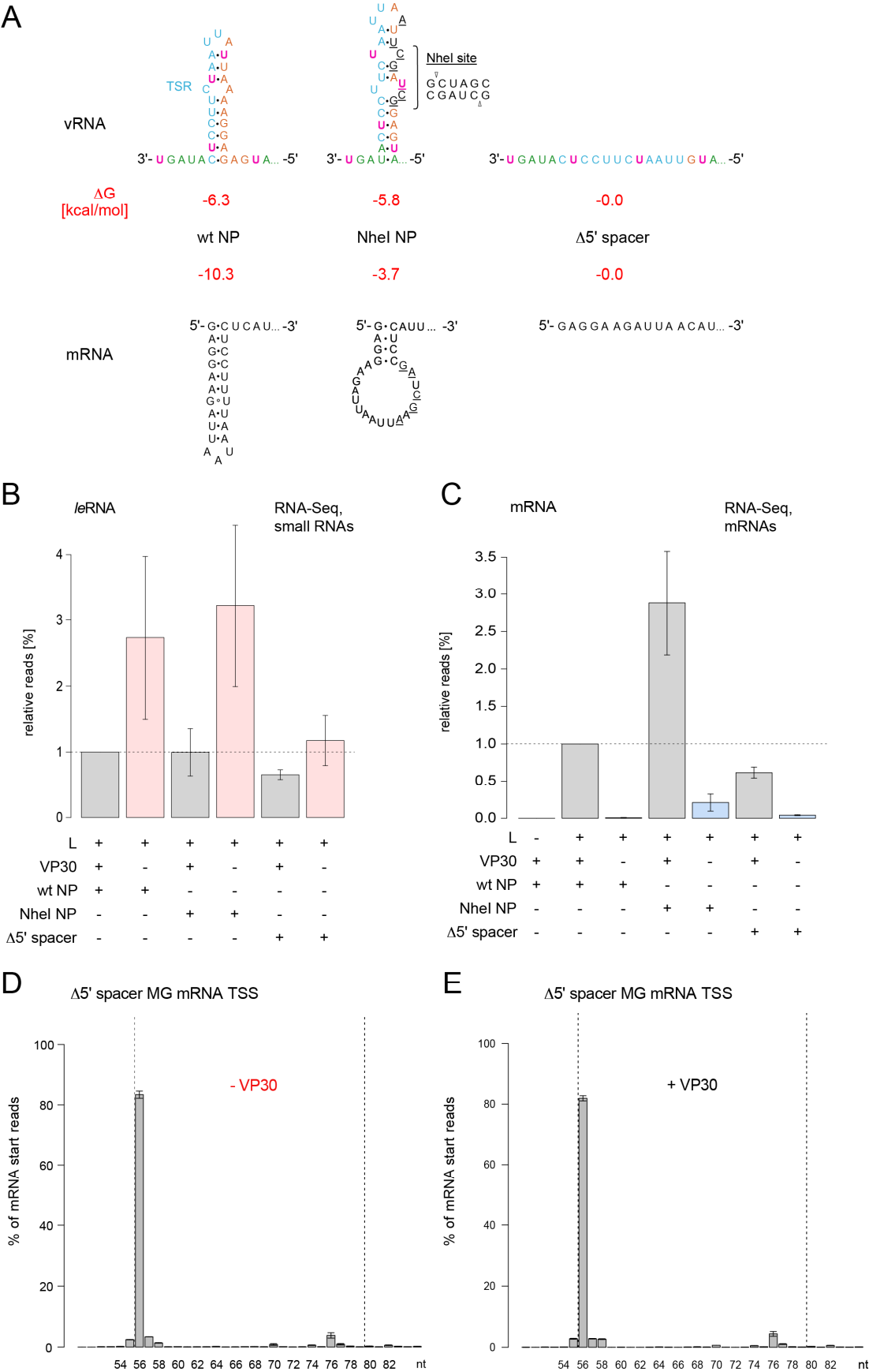


Fig. 8

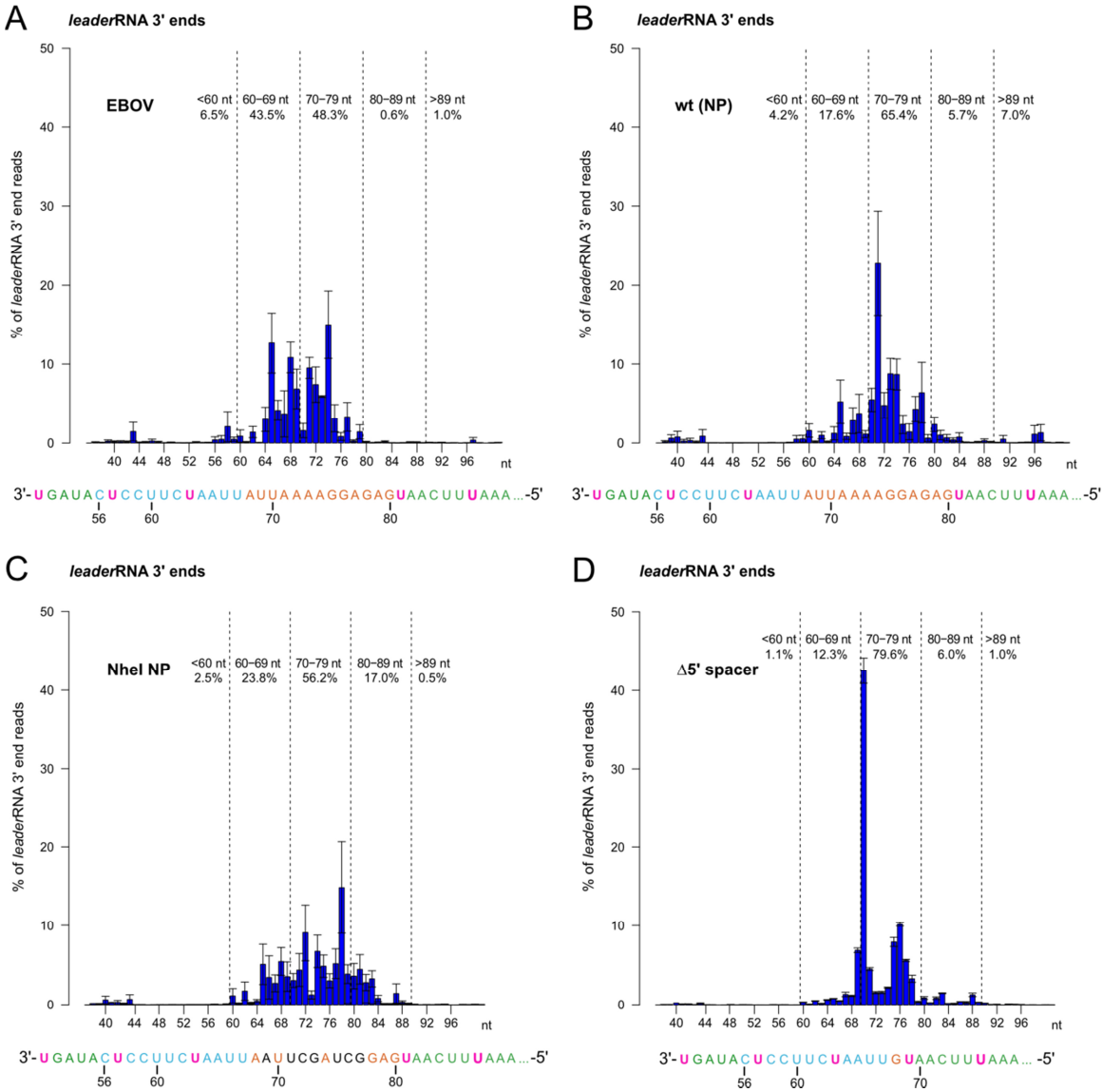


Fig. 9

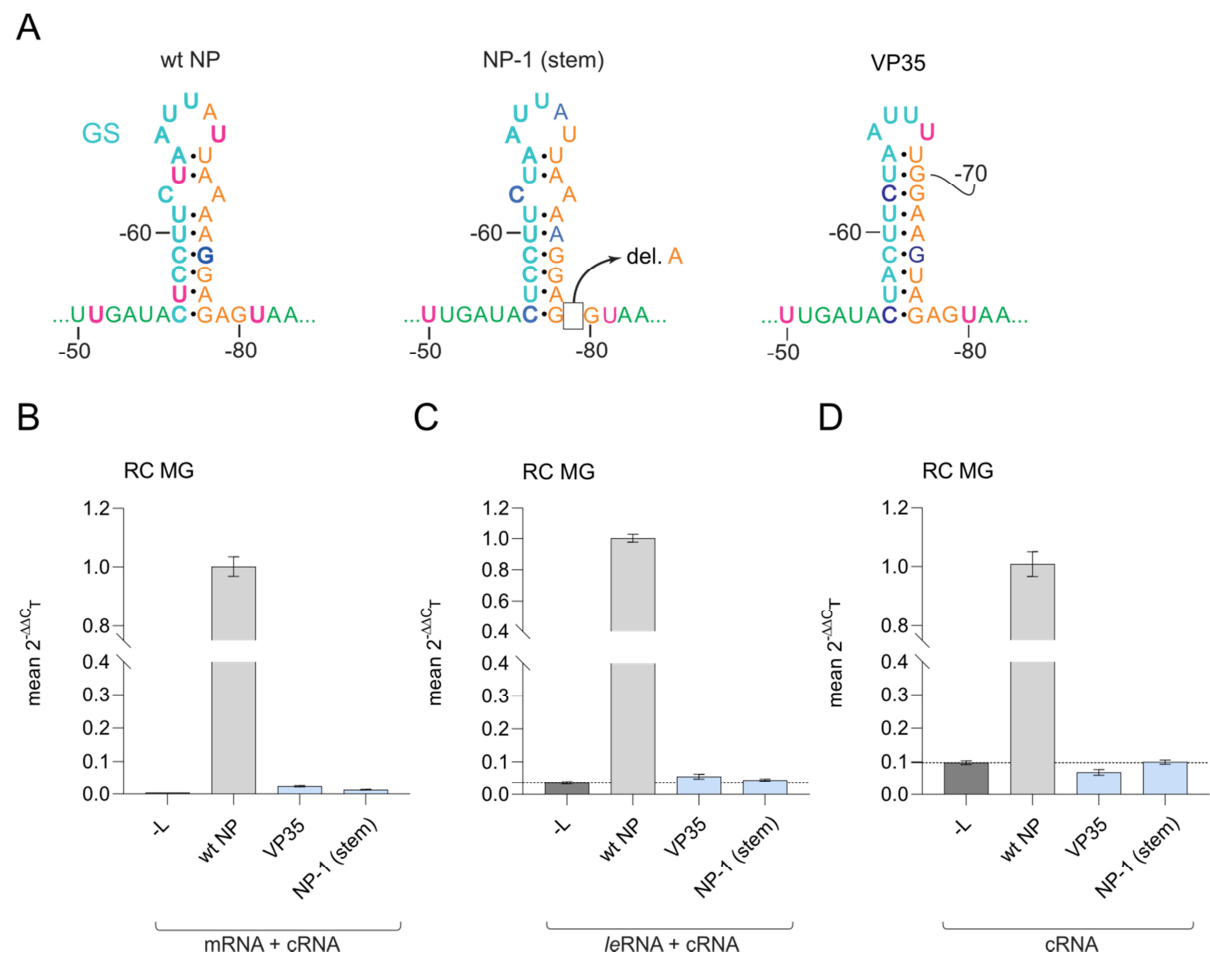


Fig. 10

## **SUPPLEMENTARY**

### **Ebola virus transcriptome analysis reveals insights into transcriptional regulation**

Simone Bach<sup>1</sup>, Julia Schlereth<sup>1</sup>, Nadine Biedenkopf<sup>2</sup>, Stephan Becker<sup>2</sup>, Marcus Lechner<sup>3</sup>, Roland K. Hartmann<sup>1,#</sup>

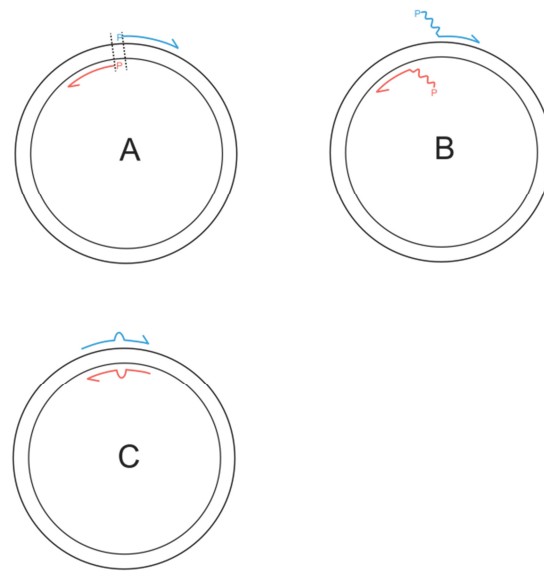
<sup>1</sup> Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35037 Marburg, Germany

<sup>2</sup> Institut für Virologie, Philipps-Universität Marburg, Hans-Meerwein-Str. 2, 35043 Marburg

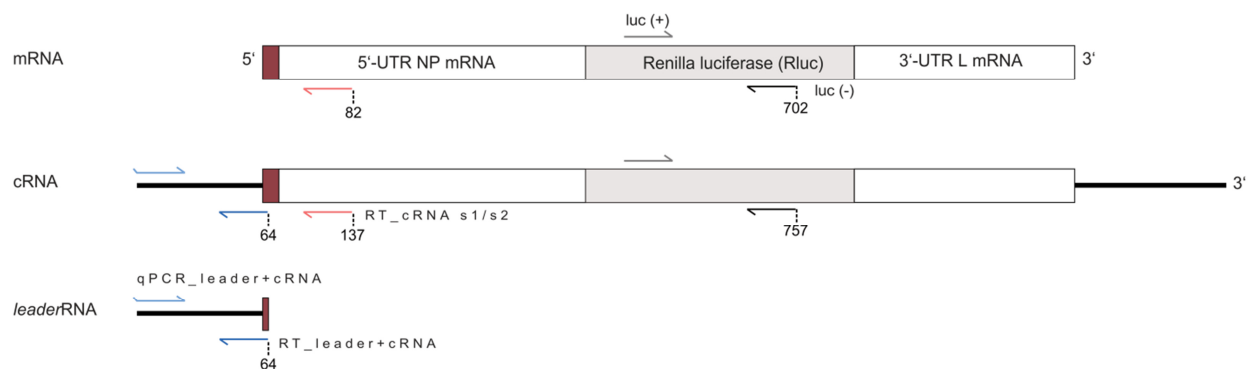
<sup>3</sup> Center for Synthetic Microbiology, Hans-Meerwein-Str. 6, 35032 Marburg



**Fig. S1:** PCR-based strategies (essentially performed as previously described, Bach et al., 2019a) that were used for the construction of mutant minigenomes. **(A)** Inside-out primer deletion mutagenesis. **(B)** Overhang/inside-out primer insertion mutagenesis. **(C)** Complementary primer mutagenesis for introduction of insertions or substitutions. In approaches A and B the entire plasmid is amplified with 5'-phosphorylated primers that introduce the desired insertions/deletions, followed by circularization of PCR products and template removal by *DpnI* treatment before bacterial transformation. In approach C phosphorylation of 5'-ends and ligation are carried out by bacterial enzymes after DNA transformation.



**Fig. S2:** Scheme of qRT-PCR strategy 2. Binding sites of primers amplifying *leaderRNA*+cRNA (blue), mRNA + cRNA (gray) and cRNA alone (RT primer in red, PCR primer pair shared with *leaderRNA*+cRNA) are marked by arrows. 5'-ends of the RT primers (dashed vertical lines and nt position) are indicated. The new mRNA+cRNA primer set provides a PCR product of 112 nt which is an appropriate size compared to the other PCR products (66 nt).



**Fig. S3:** Mathematical approach to estimate mRNA:*leader*RNA ratios in EBOV-infected HuH7 cells. An exemplary calculation for one experiments (marked in yellow in Fig. 3D of the main manuscript), based on  $\Delta C_T$  values for mRNA, cRNA and *leader*RNA+cRNA.

### Calculation of mRNA: *leader*RNA ratio

(I) calculation of  $\Delta C_T$  referring to mRNA

$$C_{T, \text{mRNA}} = 13.03; C_{T, \text{cRNA}} = 18.24; C_{T, \text{leaderRNA+cRNA}} = 15.88$$

$$\Delta C_{T(1)} = C_{T, \text{cRNA}} - C_{T, \text{mRNA}} = 18.24 - 13.03 = 5.21$$

$$\Delta C_{T(2)} = C_{T, \text{leaderRNA+cRNA}} - C_{T, \text{mRNA}} = 15.88 - 13.03 = 2.85$$

(II) calculation of  $2^{\Delta C_T}$  values

$$2^{\Delta C_{T(1)}} = 37.01$$

$$2^{\Delta C_{T(2)}} = 7.21$$

(III) estimation of mRNA vs. *leader* RNA

$$\text{mRNA} = 2^{\Delta C_{T(1)}} \times \text{cRNA} = 2^{\Delta C_{T(2)}} \times (\text{leaderRNA} + \text{cRNA})$$

$$\text{mRNA} = 37.01 \times \text{cRNA} = 7.21 \times (\text{leaderRNA} + \text{cRNA})$$

$$\curvearrowright 37.01 \times \text{cRNA} = 7.21 \times (\text{leaderRNA} + \text{cRNA})$$

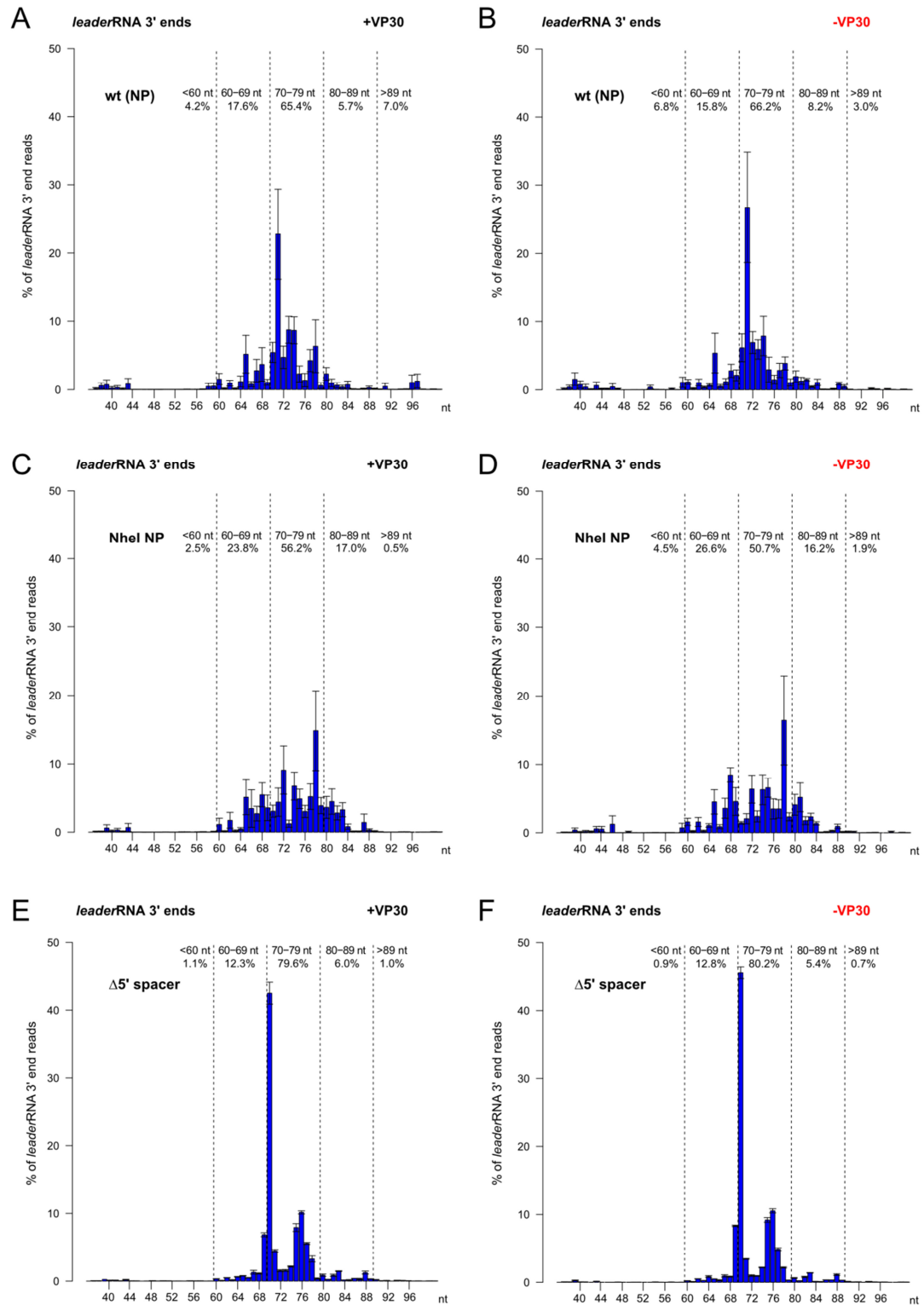
$$\curvearrowright \frac{37.01 \times \text{cRNA}}{7.21} = \text{leaderRNA} + \text{cRNA}$$

$$\curvearrowright 5.13 \times \text{cRNA} = \text{leaderRNA} + \text{cRNA}$$

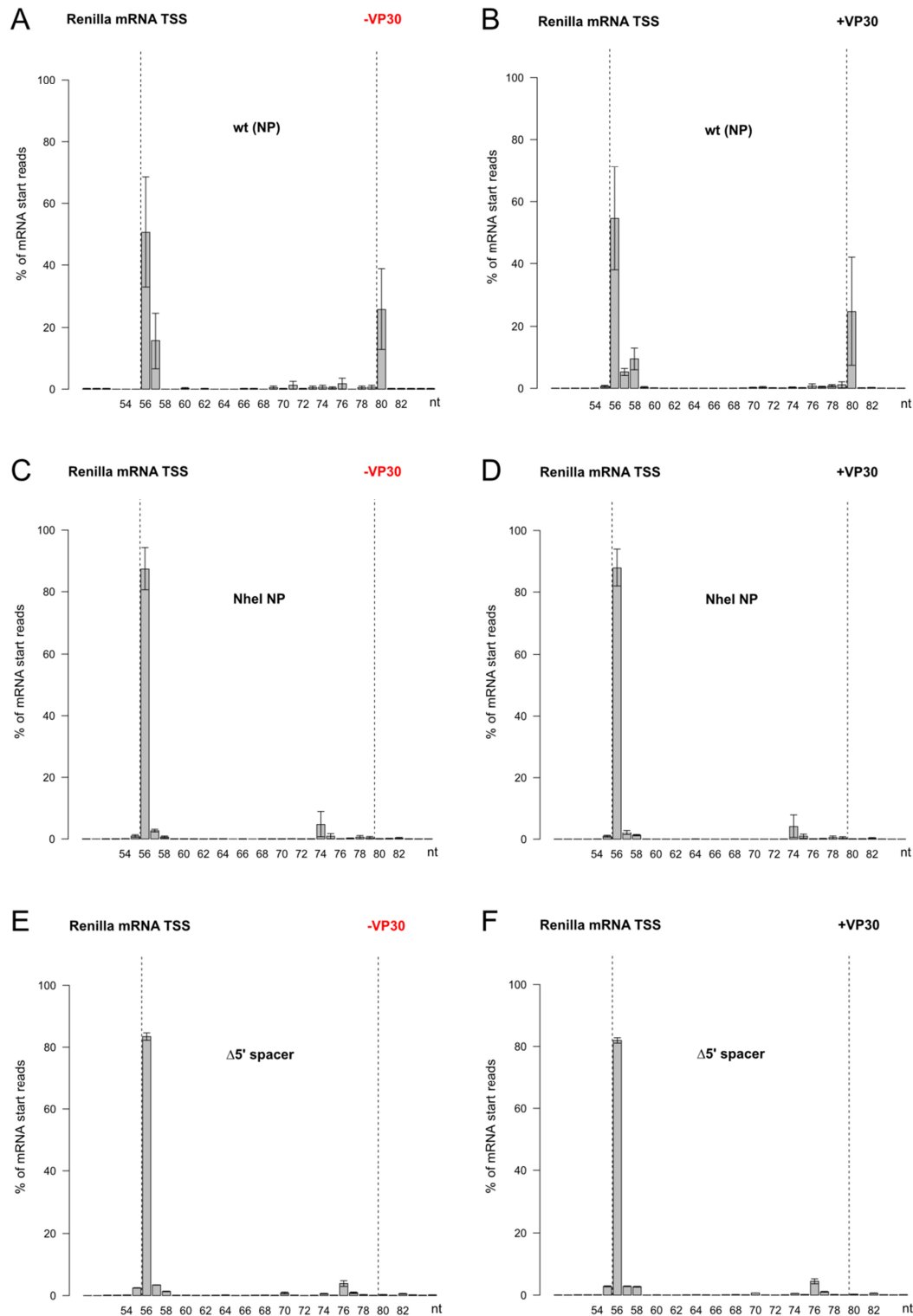
$$\curvearrowright \text{leaderRNA} = 5.13 \times \text{cRNA} - \text{cRNA} = 4.13 \times \text{cRNA}$$

$$\curvearrowright \frac{\text{mRNA}}{\text{leaderRNA}} = \frac{37.01 \times \text{cRNA}}{4.13 \times \text{cRNA}} = 8.96$$

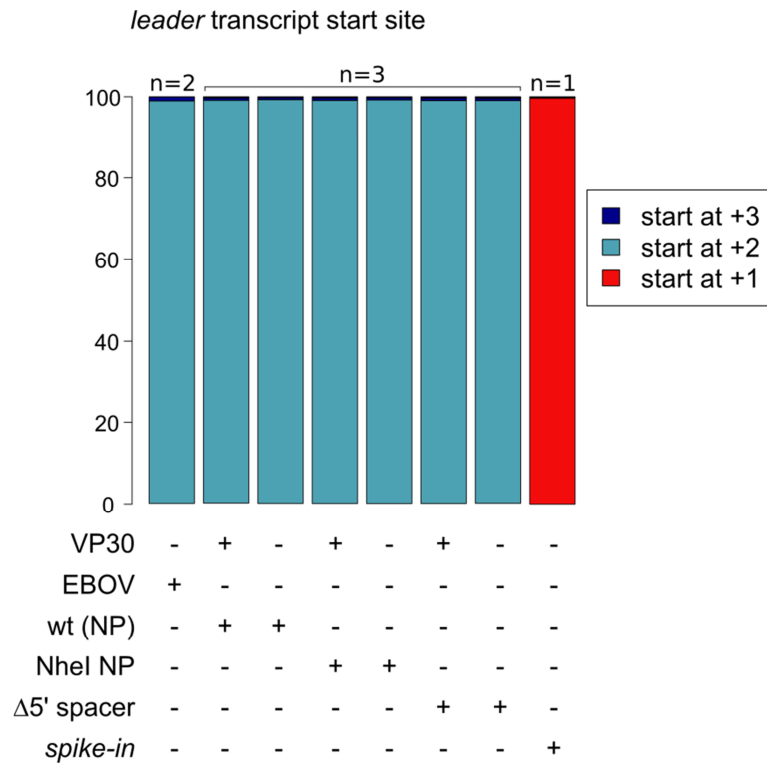
**Fig. S4:** Comparative NGS analysis of *leaderRNA* 3' termination sites in MGs (wt (NP), NheI NP,  $\Delta 5'$  spacer; illustrated in Fig. 8A of the main manuscript) in the presence (A, C, E) or absence (B, D, F) of VP30. X-axis: corresponding nt position of the respective antigenome RNA. Relative read numbers  $\pm$  SEM derive from 4 (wt MG, NheI NP MG) or 3 ( $\Delta 5'$  spacer MG) biological replicates. Relative reads accounting for certain length regions (marked by dashed vertical lines) are indicated above the bars.



**Fig. S5:** Renilla mRNA transcription start sites (TSS) of the wt (NP) MG, NheI NP MG and  $\Delta 5'$  spacer MG either in the absence (A, C, D) or presence (B, D, F) of VP30. Relative read numbers (Y-axis)  $\pm$  SEM derive from 3 independent experiments each. X-axis: nucleotide (nt) position correlated to the respective MG template RNA. Dashed vertical lines mark the corresponding genome nt -56 to -79 which harbor the TSS-spacer region separating the replication promoter elements PE1 and PE2.



**Fig. S6:** Leader transcript starting 5'-nt determined by NGS for EBOV-infected HuH7 cells (lane 1), MG-transfected HEK293 cells in the presence (+) or absence (-) of VP30 [wt (NP) MG, lanes 2-3; NheI NP MG, lanes 4-5;  $\Delta 5'$  spacer MG, lanes 6-7] and a mock (non-infected/non-transfected) HEK293 control spiked with synthetic 65-meric *leader*RNA (lane 8; sequence in table S2). For detailed information on color code see legend next to bar diagram. The number of biological replicates (n) is indicated above the respective lane.



**Table S1:** The different PCR-based strategies for site-directed mutagenesis are illustrated in Fig. S3 (except for the classical restriction cloning). The EBOV wt minigenome pANDY 3E5E is termed p3E5E\_Rluc\_RC; RC indicates the replication-competent and RD replication-deficient minigenomes. The EBOV wt bicistronic RC minigenome pA-3E5E-RLuc-FLuc is termed bici\_NP-VP35. Hyphens in primer sequences mark nucleotide deletions, boldface underlined nucleotides substitutions or insertions. In case of restriction cloning underlined nucleotides mark overhangs and boldface italic nucleotides mark restriction sites.

**Tab. 38: Cloning strategies and primers manuscript 3**

no.	name	cloning strategy	primer fwd	primer fwd (5' to 3')	primer rev	primer rev (5' to 3')
Inside-out primer deletion mutagenesis (Fig. S1 A)						
1	p3E5E_3'-Δ1	mutagenesis of p3E5E_Rluc_RC	pS_246	GGGTCGGCATGGCATCTCC	pS_245	GGACACACAAAAAGAAAG AAGAATTTTtaggATCTTT TGTG
2	p3E5E_3'-Δ2	mutagenesis of p3E5E_Rluc_RC	pS_246	GGGTCGGCATGGCATCTCC	pS_247	GACACACAAAAAGAAAGA AGAATTTTtaggATCTTT GTGTG
3	RD_p3E5E_3'-Δ1	mutagenesis of p3E5E_Rluc_RD	pS_246	GGGTCGGCATGGCATCTCC	pS_245	GGACACACAAAAAGAAAG AAGAATTTTtaggATCTTT TGTG
4	RD_p3E5E_3'-Δ2	mutagenesis of p3E5E_Rluc_RD	pS_246	GGGTCGGCATGGCATCTCC	pS_247	GACACACAAAAAGAAAGA AGAATTTTtaggATCTTT GTGTG
Cloning of pUC19_T7_80-1 for generation of Northern Blot T7-probe 80-1: PCR with pUC19_154-1 (Schlereth et al., 2016) using the primer set below. Additional G residues added to the 5'-ends of the T7 transcript for reasons of transcription efficiency are indicated in lower case and italics. For run-off T7-transcription of the probe, the resulting plasmid pUC19_80-1 was cut with <i>EcoRI</i> .						
5	pUC19_T7_80-1	see above	(-)RNA 80-1 fwd	ggGAGGAAAATTATTA ATCTTCCTCATAGTTA TTCCG	(+)RNA 56-158 rev	TATAGTGAGTCGTATTAGG ATCCTCTAGAGTC
Overhang/inside-out primer mutagenesis (Fig. S1 B)						
6	bici-NP GEmut.-leNP	mutagenesis of bici_NPGEmut.-NP	pS_189	<u>CTAAAAATTCTTCTTT</u> <u>CTTTTTGTGTGTCGGA</u> TTAGTATATACTTAATT ATAAAACGATCGTGTA AATG	pS_190	<u>GATCTTTTGTGTGCGAAT</u> <u>AACTAT</u> GAGGAAGATTAA TAATTTTCCTCATCCTTAC
Complementary primer mutagenesis (Fig. S1 C)						
7	bici_NP GEmut.-NP	mutagenesis of bici_NP-NP	pS_175	CTTCCTCATTAGT <b><u>ATA</u></b> <b><u>T</u></b> ACTTAATTATAAAAC GATCG	pS_176	CGATCGTTTTATAATTAAG <b><u>TATATA</u></b> CTAATGAGGAAG
8	bici_NP GEmut.-VP40	mutagenesis of bici_NP-VP40	pS_173	CTTCATCATTAGT <b><u>ATA</u></b> <b><u>T</u></b> ACTTAATTATAAAAC GATCGTG	pS_174	CACGATCGTTTTATAATTA AG <b><u>TATATA</u></b> CTAATGATGAA G

no.	name	cloning strategy	primer fwd	primer fwd (5' to 3')	primer rev	primer rev (5' to 3')
Complementary primer mutagenesis (Fig. S1 C)						
9	p3E5E_3' +1	mutagenesis of p3E5E_Rluc_RC	pS_248	CTTCTTTCTTTTGTG TGTCCGCGGGTCGGC ATGGCATC	pS_249	GATGCCATGCCGACCCGC GGACACACAAAAAGAAAG AAG
10	RD_p3E5E_3' +1	mutagenesis of p3E5E_Rluc_RD	pS_248	CTTCTTTCTTTTGTG TGTCCGCGGGTCGGC ATGGCATC	pS_249	GATGCCATGCCGACCCGC GGACACACAAAAAGAAAG AAG
Restriction cloning						
11	bici_NP-VP35	<b>5-step cloning:</b> 1) Backbone-PCR template: pA-3E5E-GLuc-GFP (N. Biedenkopf; AG Becker; siehe 8.2)	16	TTT <b>TTAATTAA</b> ATAGT ATCCTGATACTTGCAA AGGTTGG	17	TTT <b>CTCGAG</b> CTTGTTAGAC CAGCTTTTCTTTACATTG
		2) PCR to exchange GFP with Firefly Luc, template: pCAGGS_Luc2 (Biedenkopf and Hoenen, 2017)	18	TTT <b>CTCGAG</b> ATGGAA GATGCCAAAAACATTA AGAAGG	19	TTT <b>TTAATTAA</b> TTACACGG CGATCTTGCCG
		3) Restriction digestion of 1) and 2) with <i>PacI</i> (5'- TTAATTAA -3') and <i>XhoI</i> (5'-CTCGAG -3') followed by T4 DNA ligase ligation.				
		4) PCR to introduce NotI and NdeI restriction sites; template: p3E5E_Rluc_RC	20	TTT <b>GCGGCCGC</b> TTATT GTTCAATTTTGAGAAC TCGCTC	21	TTT <b>CATATG</b> ATGACTTCGA AAGTTTATGATCCAG
		5) Restriction digestion of 3) and 4) with <i>NotI</i> (5'- GCGGCCGC -3') and <i>NdeI</i> (5'- CATATG -3') followed by T4 DNA ligase ligation to exchange Gaussia Luc (GLuc) with Renilla Luc (RLuc).				
12	bici_NP-NP	template backbone: bici_NP-VP35; template insert: GeneArt Strings DNA fragment_NP	GeneArt Strings DNA fragment_NP (Thermo Fisher Scientific; TFS): TT <b>CTCGAG</b> CTTGTTAGACCAGCTTTTCTTTTACATTGAAGACAATAAGCTTC GAGTGCTAGAGAATTCAATTGACGTAAGGATGAGGAAAATTATTAATCTTCC TCATTAGTTTTTTCTTAATTATAAAACGATCGTGTAATGTTGGGCAGGCTAG TAGGTAAGTTATAAAAGCATAATGTAATTTAATTTAGATAAATCGCTAGTTAG AATATTGAGTTACAATTATGGCTATAAAGTGGCTTAACGCCTATTTGCATTA AGACTCAAATTGAACTTGAAAAACGGCTAGTAATAATAAGCTAGCTTTAAAT TCAAAAATAAGAATCACTTTTATTGTGATAATAATTCACACCTTAGACATCAAA AATCTTCTGTTTTCGTTCCTTGACTACTTAATGTTAGCTATTTGTCGGTTG AATCATCCCATTTGCCATGCTCAT <b>GCGGCCGC</b> GAG			
		Restriction digestion with <i>XhoI</i> (5'- CTCGAG -3') and <i>NotI</i> (5'- GCGGCCGC -3') followed by T4 DNA ligase ligation.				
13	bici_NP-VP40	template backbone: bici_NP-VP35; template insert: GeneArt Strings DNA fragment_VP40	GeneArt Strings DNA fragment_VP40 (TFS): TT <b>CTCGAG</b> CTTGTTAGACCAGCTTTTCTTTTACATTGAAGACAATAAGCTTC GAGTGCTAGAGAATTCAATTGACGTAAGGATGATGAAGGTTAATGAAAAAC ACTCTCTCAGCCGAGGTAGGTTTTCTTAATCTTCATCATTAGTTTTTTCTTAA TTATAAAACGATCGTGTAATGTTGGCAGGCTAGTAGGTAAGTTATAAAAG CATAATGTAATTTAATTTAGATAAATCGCTAGTTAGAATATTGAGTTACAATTA TGGCTATAACTGTGGCTTAACGCCTATTTGCATTAAGACTCAAATTGAACTTT GAAAAACGGCTAGTAATAATAAGCTAGCTTTAAATTCAAAAATAAGAATCACT TTTATTGTGATAATAATTCACACCTTAGACATCAAAAATCTTCTGTTTTCGT TCCTTGACTACTTAATGTAGCTATTTGTCGGTTGAATCATCCCATTTGTTCCA TGCTCAT <b>GCGGCCGC</b> GAG			
		Restriction digestion with <i>XhoI</i> and <i>NotI</i> followed by T4 DNA ligase ligation.				

**Tab. S2: RNA oligonucleotides**

no.	name	sequence	source
1	5'-PPP- <i>leader</i> RNA (65-mer <i>spike-in</i> )	pppCGG ACA CAC AAA AAG AAA GAA GAA UUU UUA GGA UCU UUU GUG UGC GAA UAA CUA UGA GGA AGA UU	AXOlabs

### RNA-Seq: cDNA library preparations and Illumina Sequencing

(A) Isolation of total RNA from EBOV-infected HuH7 cells (experiment no. 1439, 1647) or from MG-transfected HEK293 cells (experiment no. 1484, 1647, 1842). Poly(A) +RNA was isolated from total RNA. First-strand cDNA synthesis was performed using N<sub>6</sub> randomized primers, followed by RNA fragmentation with ultrasound (2 pulses of 30 s at 4°C). Then, the Illumina TruSeq sequencing adapters were ligated in a strand-specific manner to the 5' and 3' ends of the cDNA. Sample-specific barcodes were introduced by 5' adapters only (1439, 1484, 1647, 1757) or by both, 3' and 5' adapters (1842). The cDNA was finally amplified with PCR (12 or 14 cycles) using a high fidelity DNA polymerase. For Illumina sequencing, the cDNA was purified (desalted) using the Agencourt AMPure XP kit (Beckman Coulter) and fractionated, that is, cDNA products in the size range of 400 – 600 bp (exp. 1439, 1484), 350 – 700 bp (exp. no. 1647) or 300 – 500 bp (exp. no. 1842) were eluted from preparative agarose gels.

B) Isolation of small RNA from EBOV-infected HuH7 or from MG-transfected HEK293 cells: exp. no. 1439, EBOV-infected HuH7 cells: small RNAs (< 200 nt) were extracted and enriched using the mirVana™ miRNA Isolation Kit (with phenol; TFS) according to manufacturer's protocol; exp. no. 1484, MG-transfected HEK293 cells: small RNAs (< 200 nt) extracted and enriched using the mirVana™ miRNA Isolation Kit (see above); exp. no. 1647 and 1757, small RNA fractions (< 200 nt) from EBOV-infected HuH7 cells were isolated from total RNA preparations using the RNeasy MinElute Cleanup kit (Qiagen, exp. no. 1647) or RNeasy Plus Mini Kit (Qiagen, exp. no. 1757). Small RNAs (< 200 nt) from MG-transfected HEK293 cells were extracted and enriched using the mirVana™ miRNA Isolation Kit (see above); exp. no. 1842, small RNAs (< 200 nt) from MG-transfected HEK293 cells were extracted and enriched using the mirVana™ miRNA Isolation Kit (see above).

For exp. no. 1439, 1484 and 1647, ribosomal RNA molecules were depleted by using the Ribo-Zero rRNA (Gold) Removal Kit for Human-Mouse-Rat (Epicentre). The sample was split into two halves and both half samples were poly(A)-tailed using poly(A) polymerase. For exp. no. 1439, one half sample was then treated with Terminator exonuclease



(+TEX, Epicentre) to enrich for RNAs carrying a 5'-cap or a 5'-triphosphate end, while the other half sample was left untreated. For exp. no. 1484 and 1647, samples were not split into two halves after rRNA depletion, and entire samples were treated with TEX. Then, 5'-cap structures and 5'-triphosphates were converted to 5'-monophosphates using tobacco acid pyrophosphatase (TAP; exp. no. 1439) or CAP-Clip Acid Pyrophosphatase (Cellsript; exp. no. 1647). Afterwards, an RNA adapter was ligated to the 5'-monophosphate of RNAs. First-strand cDNA synthesis was performed using an oligo(dT)-adapter primer and the M-MLV reverse transcriptase. The resulting cDNAs were PCR-amplified (12-15 cycles) using a high fidelity DNA polymerase. The cDNAs were fractionated and purified using the Agencourt AMPure XP kit (Beckman Coulter). Fractionation was bead-based and provided – depending on the beads used – cDNA length ranges of either 140 – 350 bp (1439, 1484) or 200 – 500 bp (1647). TruSeq sequencing was performed using an Illumina MiSeq System with either a paired-end read-length of 2x 300 nt (1439, 1484) or a single read-length of 150 nt (1647).

For exp. no. 1757, rRNA depletion was omitted. MG RNA samples of RNA-seq 1647 were re-analyzed (Fresh library preparation). All samples were treated with TEX and CAP-Clip Acid Pyrophosphatase. Then, RNAs from EBOV-infected and MG-transfected cells were then additionally fractionated to < 100 nt by preparative 10 % denaturing PAGE and passive gel elution, ethanol precipitation and re-dissolving in water. Subsequently, oligonucleotide adapters were ligated to the 5' and 3' ends of these RNAs. 5' sequencing adapters harbored barcode sequences that were specific for each sample. First-strand cDNA synthesis was performed using M-MLV reverse transcriptase and the 3' adapter as primer. The resulting cDNAs were amplified with PCR (16-17 cycles) using a high fidelity DNA polymerase. The cDNA was purified using the Agencourt AMPure XP kit (Beckman Coulter) and analyzed by capillary electrophoresis (cDNA length: ~ 150 – 220 bp). Sequencing was performed on an Illumina NextSeq 500 system using 150-nt read length.

For exp. no. 1842, RNA samples were treated with CAP-Clip Acid Pyrophosphatase to generate 5'-monophosphate ends. Then rRNA was depleted using the Ribo-Zero rRNA Removal Kit GOLD for Human-Mouse-Rat (Illumina). Oligonucleotide adapters were ligated to the 5' and 3' ends. Both, 3' and 5' adapters, harbored TruSeq barcode sequences that were specific for each sample. First-strand cDNA synthesis was performed using M-MLV reverse transcriptase and the 3' adapter as primer. The resulting cDNAs were PCR-amplified (12 cycles) using a high fidelity DNA polymerase. The cDNA was purified using the Agencourt AMPure XP kit (Beckman Coulter) and analyzed by capillary electrophoresis. The cDNA samples were pooled and size-

fractionated to 140 – 270 bp by preparative agarose gel electrophoresis. Fractionated cDNAs were pooled with the poly(A) RNA-derived cDNA pool of this sequencing experiment (see (A), exp. no 1842). Sequencing was performed on an Illumina NextSeq 500 system using 150 bp read length.

### 1893 – spike-in control

Small RNA ribosomal RNA molecules were depleted, using the Ribo-Zero rRNA Removal Kit GOLD for Human-Mouse-Rat (Illumina). 2% of a synthetic 65-meric *leaderRNA* was added (*spike-in*) to a mock small RNA control sample. Afterwards, CAP structures and 5'-triphosphate ends were removed from the small RNA sample, using CAP-Clip Acid Pyrophosphatase (Cellscript). Oligonucleotide adapters were ligated to the 5' and 3' ends of the CAP-Clip Acid Pyrophosphatase treated and rRNA-depleted small RNA sample. First-strand cDNA synthesis was performed using M-MLV reverse transcriptase and the 3' adapter as primer. The resulting cDNA was PCR-amplified using a high fidelity DNA polymerase (12 cycles). The cDNA was purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics). The cDNA sample was size fractionated in the size range of 160 – 500 bp using a preparative agarose gel. Sequencing was performed on an Illumina NextSeq 500 system using 75 bp read length.

### RNA-Seq Analysis

RNA-Seq reads were trimmed using cutadapt (Martin, 2011) in three iterations removing the given 5'- and 3'-adapters and poly(A)-tails with a quality cutoff of 28. For NextSeq libraries, NextSeq-Trimming was used which also removes 3' G's from the sequences (equals high quality black which might or might not be a base). Minimal length of the remaining sequence had to be 16 nt. Here, no more than 10 % N's (unclear base identity) were allowed. The trimmed sequences were mapped to the Human Transcriptome GRCh38 (NCBI, Genome Reference Consortium Human Build 38 patch release 13) and, according to the respective sample, the wt (NP) Ebola MG as well as the Nhel NP- and  $\Delta 5'$  spacer-minigenome sequences or the Ebola virus genome (NCBI, 1976 strain Mayinga / NC\_002549.1) using segemehl v0.2 (Hoffmann et al., 2009) with an e-value threshold of  $10^{-8}$ .

sRNA and mRNA counts and statistics were derived using Perl-scripts. Table S3 gives an overview of all samples. A read was accounted as leader read if it started at nucleotide 1 to 5 relative to the genome of the Ebola virus. A read that started between position 1 and 106 and ended beyond position 90 was accounted as NP mRNA (this focusses on transcription starts rather than internal mRNA fragments). Note that while

listed in the Table S3, statistical relevant NP mRNA reads were only used from total RNA libraries while leader reads were used from small RNA libraries. For relative quantification of reads obtained for MG mutants +/-VP30, the read counts were divided by the counts of corresponding wt NP MG +VP30 samples. Figures were plotted using R (The R Development Core Team, 2011).

**Tab. S3:** MG, minigenome; mVP30, minus VP30; pVP30, plus VP30; QC, Quality Control; Reads mapped: % values in parentheses: reads that could be mapped to the human transcriptome and EBOV genome/antigenome; NP mRNA reads and leader reads: % values in parentheses: relative mapped reads.

**Tab. S3**

Experiment/ Replicate	Total reads (after QC)	Reads mapped	NP mRNA reads	leaderRNA reads	Experiment no.
HuH7-mock-smallRNA-1	1008615	893247 (88.5%)	0 (0%)	0 (0%)	1647
HuH7-mock-smallRNA-2	19907003	14773787 (74.2%)	1 (0%)	4 (0%)	1757
HuH7-mock-totalRNA-1	917585	515388 (56.1%)	2 (0%)	0 (0%)	1647
HuH7-mock-totalRNA-2	2360734	1476248 (62.5%)	2 (0%)	0 (0%)	1757
HuH7-EBOV smallRNA-1	3113676	2748560 (88.2%)	98 (0%)	2269 (0.08%)	1439
HuH7-EBOV smallRNA-2	1125380	1007550 (89.5%)	0 (0%)	17 (0%)	1647
HuH7-EBOV smallRNA-3	22223516	16231822 (73%)	235 (0%)	12786 (0.07%)	1757
HuH7-EBOV totalRNA-1	6399980	3732096 (58.3%)	257507 (6.89%)	190 (0%)	1439
HuH7-EBOV totalRNA-2	902477	499811 (55.3%)	2087 (0.41%)	0 (0%)	1647
HuH7-EBOV totalRNA-3	2614739	1601165 (61.2%)	5592 (0.34%)	42 (0%)	1757
MG-NheI-mVP30-smallRNA-1	4899300	4809065 (98.1%)	29 (0%)	179 (0%)	1482
MG-NheI-mVP30-smallRNA-2	870424	771580 (88.6%)	6 (0%)	49 (0%)	1647
MG-NheI-mVP30-smallRNA-3	21198913	10984500 (51.8%)	4602 (0.04%)	115289 (1.04%)	1757
MG-NheI-mVP30-smallRNA-4	15945623	11911562 (74.7%)	5024 (0.04%)	58588 (0.49%)	1842
MG-NheI-mVP30-smallRNA-5	11061387	7779682 (70.3%)	1674 (0.02%)	27993 (0.35%)	1842
MG-NheI-mVP30-totalRNA-1	8085543	5436217 (67.2%)	23057 (0.42%)	19 (0%)	1482

Experiment/ Replicate	Total reads (after QC)	Reads mapped	NP mRNA reads	leaderRNA reads	Experiment no.
MG-NheI-mVP30- totalRNA-2	1440452	906533 (62.9%)	726 (0.08%)	0 (0%)	1647
MG-NheI-mVP30- totalRNA-3	3231096	2122490 (65.6%)	1052 (0.04%)	0 (0%)	1757
MG-NheI-mVP30- totalRNA-4	16742620	9073403 (54.1%)	7852 (0.08%)	2 (0%)	1842
MG-NheI-pVP30- smallRNA-1	4788791	4744247 (99%)	5 (0%)	16 (0%)	1482
MG-NheI-pVP30- smallRNA-2	879809	779228 (88.5%)	3 (0%)	28 (0%)	1647
MG-NheI-pVP30- smallRNA-3	17587649	8476719 (48.1%)	864 (0.01%)	22565 (0.26%)	1757
MG-NheI-pVP30- smallRNA-4	17245502	13135996 (76.1%)	1796 (0.01%)	32139 (0.24%)	1842
MG-NheI-pVP30- smallRNA-5	12033890	8782594 (72.9%)	1562 (0.01%)	17906 (0.2%)	1842
MG-NheI-pVP30- totalRNA-1	8480111	5956651 (70.2%)	194521 (3.26%)	31 (0%)	1482
MG-NheI-pVP30- totalRNA-2	1305741	803329 (61.5%)	20035 (2.49%)	0 (0%)	1647
MG-NheI-pVP30- totalRNA-3	3548211	2232480 (62.9%)	34792 (1.55%)	0 (0%)	1757
MG-NheI-pVP30- totalRNA-4	16531516	9641065 (58.3%)	103504 (1.07%)	4 (0%)	1842
MG-Δ5' spacer- mVP30-smallRNA-1	13829807	10821705 (78.2%)	11765 (0.1%)	18790 (0.17%)	1842
MG-Δ5' spacer- mVP30-smallRNA-2	10902772	7704542 (70.6%)	11316 (0.14%)	23615 (0.3%)	1842
MG-Δ5' spacer- mVP30-smallRNA-3	8116776	5966700 (73.5%)	6342 (0.1%)	11872 (0.19%)	1842
MG-Δ5' spacer- mVP30-totalRNA-1	15813896	8669527 (54.8%)	2838 (0.03%)	4 (0%)	1842
MG-Δ5' spacer- mVP30-totalRNA-2	17963625	10754811 (59.8%)	1716 (0.01%)	6 (0%)	1842
MG-Δ5' spacer- mVP30-totalRNA-3	18519122	10584970 (57.1%)	3715 (0.03%)	9 (0%)	1842
MG-Δ5' spacer- pVP30-smallRNA-1	11909591	9242561 (77.6%)	3977 (0.04%)	8871 (0.09%)	1842
MG-Δ5' spacer- pVP30-smallRNA-2	15380787	11524155 (74.9%)	5472 (0.04%)	17294 (0.15%)	1842
MG-Δ5' spacer- pVP30-smallRNA-3	7556633	5328108 (70.5%)	2337 (0.04%)	5585 (0.1%)	1842
MG-Δ5' spacer- pVP30-totalRNA-1	16498553	9402355 (56.9%)	54645 (0.58%)	8 (0%)	1842
MG-Δ5' spacer- pVP30-totalRNA-2	15723665	9387342 (59.7%)	44318 (0.47%)	6 (0%)	1842

Experiment/ Replicate	Total reads (after QC)	Reads mapped	NP mRNA reads	leaderRNA reads	Experiment no.
MG-Δ5' spacer- pVP30-totalRNA-3	14082457	8269347 (58.7%)	42995 (0.51%)	4 (0%)	1842
MG-spikein- smallRNA-1	9022658	8122324 (90%)	0 (0%)	6027028 (74.2%)	1893
MG-wt-mL- totalRNA-0	16707013	9806296 (58.6%)	46 (0%)	2 (0%)	1842
MG-wt-mVP30- smallRNA-1	4340919	4271420 (98.3%)	23 (0%)	175 (0%)	1482
MG-wt-mVP30- smallRNA-2	1236716	1074289 (86.8%)	2 (0%)	71 (0%)	1647
MG-wt-mVP30- smallRNA-3	19254766	11337403 (58.8%)	10588 (0.09%)	61445 (0.54%)	1757
MG-wt-mVP30- smallRNA-4	26537720	22142500 (83.4%)	4411 (0.01%)	46169 (0.2%)	1842
MG-wt-mVP30- smallRNA-5	7138030	5137782 (71.9%)	1300 (0.02%)	17307 (0.33%)	1842
MG-wt-mVP30- totalRNA-1	6028736	4427489 (73.4%)	868 (0.01%)	8 (0%)	1482
MG-wt-mVP30- totalRNA-2	1163508	714209 (61.3%)	7 (0%)	0 (0%)	1647
MG-wt-mVP30- totalRNA-3	2828757	1785376 (63.1%)	12 (0%)	0 (0%)	1757
MG-wt-mVP30- totalRNA-4	16163017	8624088 (53.3%)	372 (0%)	5 (0%)	1842
MG-wt-pVP30- smallRNA-1	4112460	4084837 (99.3%)	8 (0%)	23 (0%)	1482
MG-wt-pVP30- smallRNA-2	825900	729879 (88.3%)	2 (0%)	55 (0%)	1647
MG-wt-pVP30- smallRNA-3	18737682	11409156 (60.8%)	2171 (0.01%)	64823 (0.56%)	1757
MG-wt-pVP30- smallRNA-4	12891822	9492050 (73.6%)	2540 (0.02%)	29733 (0.31%)	1842
MG-wt-pVP30- smallRNA-5	6839812	3399475 (49.7%)	256 (0%)	7640 (0.22%)	1842
MG-wt-pVP30- totalRNA-1	4632723	3433850 (74.1%)	41673 (1.21%)	7 (0%)	1482
MG-wt-pVP30- totalRNA-2	1194709	754977 (63.1%)	6644 (0.88%)	1 (0%)	1647
MG-wt-pVP30- totalRNA-3	3454961	2324874 (67.2%)	13724 (0.59%)	4 (0%)	1757
MG-wt-pVP30- totalRNA-4	18244781	10163327 (55.7%)	79174 (0.77%)	38 (0%)	1842

#### **4.4 Der Einfluss von RNA-Sequenz, -Struktur und Länge der TSS-Spacer-Region auf die VP30-Abhängigkeit der Transkription und die Balance von Transkription und Replikation**

Die virale Transkription des Ebola-Virus (EBOV) wird vom Transkriptionsfaktor VP30 reguliert. Die Aktivität von VP30 beruht auf seinen RNA-Bindeeigenschaften, seiner dynamischen Phosphorylierung sowie seiner Interaktion mit den Nukleokapsid-Proteinen VP35 und NP. Es gibt Hinweise, dass eine RNA-Hairpin-Struktur am Transkriptionsstartsignal (TSS) des ersten NP-Gens (NP-Hairpin) die VP30-Abhängigkeit der viralen Transkriptionsinitiation vermittelt. In der vorliegenden Studie haben wir die Anforderungen an die RNA-Sequenz und -Struktur der Transkriptionsstartregion für die VP30-abhängige Transkription untersucht. Hierfür wurden Replikations-kompetente und -defiziente Minigenomsysteme verwendet, in denen der NP-Hairpin entweder eliminiert, mutiert oder durch Genstart-Hairpins interner EBOV-Gene, bzw. Mutationsvarianten davon, ersetzt wurde. Die Ergebnisse zeigen auf, dass Abweichungen von Sequenz, Struktur oder Länge des nativen NP-Hairpins, die mit einer Destabilisierung des Hairpins einhergehen, die stringente VP30-Abhängigkeit der viralen Transkription am TSS partiell, jedoch je nach Variante unterschiedlich stark relaxieren. Dagegen führen Stabilisierungen der Haarnadelstruktur zu einer stringenteren Kontrolle der Transkription durch VP30. Eine G<sub>75</sub> zu U-Punktmutation, die ein kontinuierliches 3'-UN<sub>5</sub>-Hexamerraster zwischen den Promotorelementen PE1 und PE2 etabliert, führte zur höchsten Rate an VP30-unabhängiger Transkription. Diese blieb von der Sekundärstrukturstabilität am TSS weitgehend unbeeinträchtigt. Darüber hinaus wurde die Transkription selbst durch eine Deletion des Spacer-Bereichs (12 Nt) zwischen PE1 und PE1, die mit der Elimination jeglicher RNA-Sekundärstruktur verbunden war, nicht substanziell beeinträchtigt. Hairpin-Stabilisierungen und die U<sub>75</sub>-Mutation eliminierten zudem den relativen Anstieg der Replikation in Abwesenheit von VP30. Zusammenfassend deuten die experimentellen Daten darauf hin, dass sowohl Sequenz (-Länge) als auch die spezifische Struktur des nativen NP-Hairpins im PE1-PE2-Spacerbereich evolutionär optimiert wurden, (i) zum einen, um eine stringente Transkriptionsregulation durch VP30 zu gewährleisten bei gleichzeitiger Vermeidung zu robuster Sekundärstrukturen, die die Aktivität der viralen Polymerase allgemein beeinträchtigen, (ii) zum anderen, um in Abwesenheit von VP30 bzw. nach Dissoziation des Faktors vom Polymerasekomplex einen Wechsel hin zu einer verstärkten Replikationsaktivität der Polymerase zu ermöglichen.

**Sequence, structure and length constraints in the spacer region of the genomic EBOV promoter on VP30 dependency of transcription and the balance of transcription and replication**

Simone Bach<sup>1</sup>, Nadine Biedenkopf<sup>2</sup>, Stephan Becker<sup>2</sup>, Roland K. Hartmann<sup>1,#</sup>

<sup>1</sup>Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35037 Marburg, Germany

<sup>2</sup>Institut für Virologie, Philipps-Universität Marburg, Hans-Meerwein-Str. 2, 35043 Marburg

#corresponding author:

E-mail:       roland.hartmann@staff.uni-marburg.de

---

**ABSTRACT**

Viral transcription and replication of Ebola virus (EBOV) is balanced by the transcription factor VP30. The protein's activity was shown to be dependent on RNA binding, dynamic phosphorylation and interaction with the nucleocapsid proteins VP35 and NP. A hairpin structure at the transcription start signal (TSS) of the first NP gene is thought to mediate the VP30 dependency of transcription. Here, we comprehensively studied the constraints of VP30-dependent transcription at the TSS using a variety of replication-competent and -deficient monocistronic minigenomes, in which the NP hairpin was either eliminated, mutated or replaced with hairpin structures of gene start (GS) regions of internal EBOV genes and mutated variants thereof. Sequence, structure and length deviations from the native NP hairpin led to changes in the extent of VP30-independent transcription in all tested variants. Hairpin destabilizations tended to relax VP30 dependency, while stabilizations increased VP30 dependency. A G<sub>75</sub> to U point mutation, introducing a continuous 3'-UN<sub>5</sub> hexamer phasing between replication promoter elements PE1 and PE2, led to the highest extent of VP30-independent transcription, essentially unaffected by overall hairpin stability. Furthermore, deletion of 12 nt between PE1 and PE2, which abolished any RNA secondary structure formation potential, did not eliminate the basic requirement of VP30 for efficient transcription. Hairpin stabilizations and the U<sub>75</sub> mutation also abolished the relative increase of replication in the absence of VP30. Overall, our results suggest that sequence (length) and structure of the native NP hairpin in the spacer between PE1 and PE2 are optimized (i) for tight regulation by VP30 while avoiding an extent of hairpin stability that impairs viral transcription, (ii) for enabling the switch from transcription to replication when VP30 is not part of the polymerase complex.



## INTRODUCTION

Ebola virus (EBOV) is a highly pathogenic member of the family of *filoviridae* in the order *Mononegavirales* with a high potential to cause severe hemorrhagic fever in humans with fatality rates of up to 90 % (Rougeron et al., 2015). Its non-segmented negative sense (NNS) RNA genome is encapsidated by the nucleoprotein NP and serves as template for viral replication and transcription of 9 individual viral mRNAs. Both processes are catalyzed by the viral RNA-dependent RNA polymerase L and its cofactor VP35 (abbreviated as RdRp). Replication of (-) RNA genomes (viral RNA, vRNA) requires a 2-step mechanism, that is, synthesis of an intermediate positive sense antigenomic/copy RNA (cRNA) that serves as the template for the synthesis of new RNA genomes (Mühlberger, 2007). cRNA synthesis is initiated opposite genome position -2 (Deflubé et al., 2019). EBOV transcription is sequential and follows a stop-start mechanism regulated by conserved gene start (GS) and gene end (GE) sequences. Sequential stop-start transcription of mRNAs and a single entry site of the RdRp complex at the genome 3' end result in an mRNA gradient (Fig. 1A) with declining mRNA amounts from the first NP gene to the last (L) gene (Albariño et al., 2018; Mühlberger, 2007; Shabman et al., 2014). Beyond the RdRp components L and VP35, viral transcription further requires the transcription activator VP30 (Feldmann et al., 1992; Mühlberger et al., 1999; Sanchez and Kiley, 1987). The transcription support function of VP30 is dependent on its capability to bind RNA, on a dsRNA-mediated interaction with VP35 and its phosphorylation status (Biedenkopf et al., 2016b, 2016a; Schlereth et al., 2016). Non-phosphorylated VP30 boosts transcription and simultaneously suppresses replication. Phosphorylation of VP30, particularly at Ser<sub>29</sub>, increases the binding affinity for NP and decreases its affinity for VP35. These changes in binding affinities are thought to trigger the release of VP30 from the transcription complex, resulting in a shutdown of transcription and enhancement of viral replication (Biedenkopf et al., 2013, 2016a). Recently, VP35 phosphorylation at Thr<sub>210</sub> was shown to be a key feature for the transcriptase function of the EBOV RdRp and was correlated with improved VP35:NP binding (Ivanov et al., 2019). Based on an *in vitro* study, the RNA binding protein VP30 preferentially binds ssRNA of mixed base composition in conjunction with an RNA secondary structure (Schlereth et al., 2016). This is in line with another study using EBOV-specific minigenome (MG) systems that proposed an influence of RNA secondary structure formation at the first NP GS signal (NP hairpin structure, Fig. 1B) on VP30-dependent RNA synthesis (Weik et al., 2002). This was inferred from the observation that transcription of MGs comprising a wild-type (wt) NP hairpin (HP) structure only occurred in the presence of VP30, while a mutant MG with a destabilized hairpin

(NheI NP; Weik et al., 2002) had a relaxed VP30 requirement. Furthermore, a recent study reported that the NP hairpin influences the interaction between NP and VP30 (Xu et al., 2017). It is noteworthy that the NP HP at the TSS, plus two additional nucleotides, separate the two elements of the EBOV bipartite replication promoter (PE1 and PE2; Fig. 1B). It has been shown that hexamer phasing ("rule of 6") between PE1 and PE2 is key to productive replication and transcription (Weik et al., 2005; Bach et al., 2019a). Of note, RNA secondary structures are predicted as well for internal GS regions, both on the genomic and antigenomic strand (Fig. 1A; Mühlberger et al., 1996; Sanchez et al., 1993). The structures are likely to form in the viral mRNAs, but it is unclear if they may also form (transiently) in the genomic RNA or during antigenome synthesis. Further evidence for the involvement of RNA secondary structure formation in viral RNA synthesis was provided by recent studies. One of these studies revealed a novel NTPase and helicase function of VP35 (Shu et al., 2019), another study showed that extensive RNA secondary structure stabilization at the TSS reduced the production of all three viral RNA species (mRNA, cRNA, and vRNA) depending on the extent of stabilization, while mRNA synthesis was affected most strongly (Bach et al., 2019b). Highly stabilized hairpin structures were able to fully eliminate RNA synthesis. Elimination of any RNA structure formation potential at the TSS by deletion of 12 nt ( $\Delta 5'$  spacer mutant) still permitted mRNA synthesis, although reduced to ~ 60 % relative to the wt NP minigenome.

In the present study we set out to define the structural requirements for the VP30 dependency of transcription in more detail. For this purpose, series of monocistronic RC and RD MGs with sequence, structure and length variations in the spacer between PE1 and PE2 were tested for reporter activity, and levels of the different viral RNA species were quantified by qRT-PCR. In all cases, reporter activity and viral RNA levels were directly compared for cells cotransfected with and without the plasmid encoding VP30. As reference for VP30-independent transcription, the previously described NheI NP and  $\Delta 5'$  spacer mutant MGs were included.

## RESULTS

### **Relaxed VP30 dependency of transcription is not simply related to weakened RNA secondary structure at the TSS**

Relaxation of VP30-dependent transcription observed for the NheI NP mutant MG has been the basis for the model that hairpin stability at the TSS (assumedly on the mRNA level) defines the control of transcription by VP30 in the EBOV system (Weik et al., 2002). We set out to scrutinize this assumption by constructing a series of minigenomes that varied in hairpin stability, length and RNA sequence in the transcriptional start region between PE1 and PE2 (Fig. 2). The replication-competent minigenomes (RC MGs), co-transfected with or without VP30, were analyzed in luciferase reporter assays (Fig. 3A). In addition to the NheI NP mutant MG as a reference for the reported VP30-independent viral transcription, we also included the  $\Delta 5'$  spacer mutant to investigate the influence of an eliminated RNA structure potential at the TSS (Bach et al., 2019b). Our initially tested set of mutants included only variants obeying the “rule of 6”, that is, 3'-UN<sub>5</sub> hexamer phasing between PE1 and PE2 shown to be mandatory for efficient replication and transcription (Weik et al., 2005; Bach et al., 2019a). All tested RC MG mutants gave rise to higher reporter activities than the wt NP MG in the absence of VP30. However, the extent of VP30-independent transcription differed considerably. Five of the nine MG constructs gave reporter activity equal to or higher than the NheI NP mutant. Surprisingly, highest VP30-independent transcription was not observed for the  $\Delta 5'$  spacer mutant lacking any RNA secondary formation potential, but for the two VP35 hairpin (VP35+1 (loop) v1 and v2) constructs. An additional C residue was inserted into their apical loop (either at genome position -69 or -68) to fulfill the rule of 6 (Bach et al., 2019a). Here, reporter activity reached 41 % and 53 %, respectively, in the absence of VP30 relative to the wt NP MG in the presence of VP30 (set to 100 %). Likewise, levels of VP30-independent transcription for these variants were ~ twofold higher than for the NheI NP and  $\Delta 5'$  spacer mutant MGs [19 % activity compared to wt NP +VP30]. The finding of higher VP30-independent transcription activity for MG variants with secondary structures (VP35+1 variants, VP40) comparably stable as that of the wt NP HP, as well as lower relaxation for the “unstructured”  $\Delta 5'$  spacer variant suggested that RNA secondary structure formation is not the only determinant of VP30 dependency. As reporter assays of RC MGs cannot disentangle effects on viral transcription and replication, we also tested several of the mutant constructs as part of a replication-deficient minigenome (RD MGs; Fig. 1B). In the RD MG context viral mRNA synthesis only takes place on nucleocapsids containing the vRNA synthesized by the T7 RNA polymerase. We basically observed the same trends for the RD MGs (Fig. 3B) as seen

for the RC MGs (Fig. 3A), although the extent of VP30-independent transcription was lower in the RD MG setup. For example, variant VP35+1 (loop) v1 gave 16 % reporter activity without VP30 (Fig. 3B) compared with ~ 41 % in the replication-competent MG context (Fig. 3A). Differences in the strength of effects observed in RD versus RC MGs may be related to the lower activities and activity range (high background) in the RD MG system, to differences in the structure of nucleocapsids and/or to different limiting steps in the two setups that may have led to overestimation of VP30 relaxation in the RC MG system or underestimation in the RD MG system.

To further address this issue, we determined the mRNA levels for selected RC MG constructs by qRT-PCR (Fig. 3C) using total RNA derived from the same cells that were used for analysis of reporter activity in Fig. 3A. qRT-PCR quantification of mRNAs confirmed the trend of VP30-independent reporter activity, although mRNA levels were lower than expected from RC MG reporter assays (Fig. 3A). This finding suggests that reporter assays using RC MGs indeed tend to overestimate the level of viral mRNA transcription. Likewise, mRNA levels in the presence of VP30 tended to be somewhat lower for the VP35+1 (loop) v1 and particularly the NheI NP mutant compared with the wt NP MG, a trend not seen in the reporter assays (cf. Fig. 3C and 3A).

### **Increased RNA secondary structure stability reduces VP30-independent transcription**

So far, we could not observe a clear correlation between VP30 dependency of transcription and RNA hairpin stability at the TSS. We addressed this issue by analyzing the NP G<sub>72</sub> mutant with a stabilized hairpin structure (Fig. 4A) and MG variants in which the NP HP was replaced with the VP40 HP and two incrementally, yet moderately stabilized versions of this hairpin (Fig. 4A). The NP G<sub>72</sub> mutant, differing from the wt NP HP by a single mutation and avoiding changes in UN<sub>5</sub> hexamer phasing, showed a further reduction of VP30-independent reporter activity relative to the wt NP MG (Fig. 4B); qRT-PCR quantification of mRNA levels then revealed that mRNA levels were decreased not only in the absence but also in the presence of VP30 relative to the respective wt NP control (Fig. 4C). Without VP30, the VP40 HP construct showed substantially increased levels of reporter activity and, to a lesser extent, increased mRNA levels (Fig. 4B, C). This phenotype tended to attenuate for the VP40<sub>6</sub> variant, which differed from the native VP40 HP by conversion of a C:U mismatch into a G:C bp in the upper stem region (Fig. 4A). The second VP40<sub>5</sub> variant with increased stem stability owing to an additional U:A bp and deletion of a bulged A residue, already showed a marked reduction in VP30-independent reporter activity and mRNA levels

(Fig. 4B). However, mRNA levels were also decreased to ~ 25 % in the presence of VP30 (Fig. 4C). We conclude from the findings presented in Fig. 5 that even minor stabilizations of hairpin structures at the TSS rapidly reduce viral transcription efficiency in the absence of VP30, but also in its presence.

### **RNA secondary structure stabilization reduces cRNA synthesis in the presence and absence of VP30**

Quantification of cRNA and vRNA levels for the MGs analyzed in Fig. 4 revealed that the relative increase of cRNA synthesis seen for the wt NP MG upon omission of VP30 was dampened in the case of the stabilized variant NP G<sub>-72</sub> (Fig. 4D). The same pertains to the VP40<sub>5</sub> and VP40<sub>6</sub> mutants relative to the VP40 construct. This trend was also evident in the relative vRNA levels (Fig. 4E). In the presence of VP30, all mutant constructs including the VP40 variant yielded lower cRNA levels than the wt NP MG (Fig. 4D).

### **The G to U<sub>-75</sub> point mutation relaxes VP30 dependency of transcription and suppresses the switch from transcription to replication in the absence of VP30**

We next addressed if 3'-UN<sub>5</sub> hexamer phasing between PE1 and PE2 may affect the VP30 dependency of transcription. A G<sub>-75</sub> to U mutation in the NP hairpin, making the UN<sub>5</sub> hexamer phasing continuous between PE1 and PE2, enhanced the efficiency of transcription and replication in the presence of VP30 (Bach et al., 2019a). This also pertains to variant NP U<sub>-75</sub>/G<sub>-72</sub> with an additional mutation (G<sub>-72</sub>) to counteract the destabilizing effect of the U<sub>-75</sub> mutation on hairpin formation on the genomic and antigenomic level (Fig. 5A). The two MG variants not only enhanced reporter activity in the RC and RD MG background in the presence of VP30, but also VP30-independent activity relative to the wt NP, NheI NP and Δ5' spacer MGs (Fig. 5B and C), consistent with increased mRNA levels determined by qRT-PCR (Fig. 4D). Remarkably, reporter activity of the NP U<sub>-75</sub>/G<sub>-72</sub> MG without VP30 exceeded that of the wt NP MG with VP30 (Fig. 5B). As discussed above, relaxation of VP30-dependency was less pronounced in the RD MG setup measuring viral transcription only (Fig. 5C). In summary, the NP U<sub>-75</sub>/G<sub>-72</sub> MG gave rise to the highest level of VP30-independent transcription measured so far (~ 34 %; Fig. 5D). Considering that reporter gene activity and mRNA levels were substantially increased for the NP U<sub>-75</sub> and NP U<sub>-75</sub>/G<sub>-72</sub> mutants, with and without VP30 and despite their difference in hairpin stability, indicates that continuous UN<sub>5</sub> hexamer phasing between PE1 and PE2 can affect VP30-dependent and

-independent transcription markedly more than RNA structure formation in the transcriptional start region.

We also determined the cRNA and vRNA levels for the same set of RC MG variants (Fig. 5E and F). The wt NP, NheI NP and  $\Delta 5'$  spacer MGs showed increased cRNA and vRNA levels in the absence of VP30, in line with a switch from transcription to replication when VP30 is absent. Remarkably, such an increase in the absence of VP30 was largely attenuated for the NP U<sub>-75</sub> and NP U<sub>-75</sub>/G<sub>-72</sub> mutants. Altogether, these mutants substantially relax VP30 dependency of transcription and weaken the switch from transcription to replication in the absence of VP30.

### **VP30 attenuates the requirement of hexamer spacing for viral transcription**

To investigate the role of hexamer phasing between PE1 and PE2 for VP30-independent transcription and replication, we quantified the mRNA, cRNA and vRNA levels for MG constructs that either obeyed hexamer phasing between genome positions -51 and -80 [wt NP; VP35+1 (loop) v1; GP-2 nt PE1] or disobeyed hexamer phasing [VP35; NP-1 (stem)]. Please note that nucleotides -54 and -55 of PE1, directly 3' of the TSS (pos. -56), are deleted in variant GP-2 nt PE1. Thus, although hexamer phasing between PE1 nt -51 and PE2 was maintained, the distance of the TSS to the genome 3' end was altered (Fig. 6A). As already shown in Fig. 3C, construct VP35+1 (loop) v1 gave rise to substantial VP30-independent mRNA synthesis compared to the wt NP MG (Fig. 6B). In the absence of VP30, mRNA levels of constructs VP35, NP-1 (stem) and GP-2 nt PE1 were essentially as low as the background (-L) control (Fig. 6B). However, in the presence of VP30, mRNA levels of the VP35 and NP-1 (stem) mutants were still low but clearly above the (-L) control. This finding suggests that low efficiency transcription of viral mRNA can occur with support of VP30 despite violation of hexamer periodicity between PE1 and PE2. However, there was not even a minor rescue of mRNA synthesis in the presence of VP30 in the case of the GP-2 nt PE1 variant. Thus, deletion of nucleotides -54 and -55 completely abolished transcription (Fig. 6B; Bach et al., 2019a).

For cRNA and vRNA synthesis (Fig. 6C, D), several aspects are worth mentioning: (i) basic levels of cRNA and vRNA synthesis above background can occur despite disruption of hexamer UN<sub>5</sub> phasing, as evident for the VP35 HP construct; (ii) this low-efficiency replication was barely affected by the presence of VP30; (iii) the GP-2 nt PE1 construct apparently still allowed the synthesis of very low cRNA levels that in turn gave rise to modest vRNA synthesis but clearly above the (-L) background; (iv) for the VP35 HP construct, cRNA levels were ~ 8-fold lower (+ VP30, Fig. 6C) than for the wt NP MG (after subtraction of -L background), but vRNA levels were only ~ 2-fold lower (+ VP30,

Fig. 6D). Evidently, vRNA synthesis is less affected than mRNA or cRNA synthesis by violation of hexamer periodicity at the genomic 3'-leader promoter. A possible explanation is that viral polymerases that failed to initiate RNA synthesis at the genomic promoter became increasingly available for replication initiation at the antigenomic promoter. Finally, the extent of residual mRNA, cRNA and vRNA synthesis in MG constructs violating the rule of 6 depends on the sequence/structure context [cf. variants VP35 and NP-1 (stem); Fig. 6B-D].

### **Spacer expansions of up to 66 nt still allow detectable VP30-independent transcription**

We previously showed that an extension of the PE1/PE2 spacer region by up to 66 nt is tolerated in the presence of VP30 in the RC MG system (Bach et al., 2019b). We thus analyzed to which extent different spacer expansion constructs supported viral transcription in the absence of VP30 (Fig. 7A). The analysis included 3'-leader variants in which the NP HP was replaced with the VP40 HP and derivatives thereof with insertions of up to 5 x 6 nt, the longest variant VP40+30 expanding the spacer region by 66 nt (Fig. 7B). Furthermore, we included a VP24 HP-derived variant (48 nt longer than the NP spacer region) and two variants carrying the gene L HP (+ 1 nt insertions to obey the rule of 6) instead of the NP HP (Fig. 7B). The relatively high level of VP30-independent reporter activity of the VP40 construct remained quite constant for variants VP40+6, +12 and +18, and was also similar to that of the VP24+4 variant with equal spacer length as the VP40+12 construct (Fig. 7A). Variants VP40+24 and +30 showed decreased activity without VP30, but also a drop in activity in the presence of VP30, suggesting that expansions of 60 and more nucleotides in the spacer between PE1 and PE2 begin to decrease the proportion of polymerases that are able to recognize the genomic promoter. A clear difference in reporter activity with and particularly without VP30 was observed for the variants L+1 (stem) versus L+1 (loop) which have the same spacer length but differ in the position of their 1-nt insertion (either in the loop or at the base of the stem) to maintain hexamer phasing. As this 1-nt insertion at the two different positions is unlikely to substantially affect the stability of the hairpin structure (Fig. 7B), we consider two possible explanations: either the difference in the positioning of one UN<sub>5</sub> hexamer [U<sub>-105</sub> in variant L+1 (stem), U<sub>-117</sub> in variant L+1 (loop); Fig. 2] is responsible for the difference, or the 1-nt insertion in the L+1 (stem) that moves the base of the genomic hairpin stem 1 nt away from PE2. A very similar pattern was observed for the VP35+1 (stem) and (loop) variants (Fig. 2A).

## DISCUSSION

The present study provides detailed insight into how VP30-dependent transcription initiation at the TSS responds to changes in RNA sequence, secondary structure formation and spacing of regulatory sequences in the genomic 3'-leader promoter. By mutational analysis and exchange of the NP hairpin structure at the TSS with putative 5'-UTR hairpin structures of internal GS regions we were able to show that initiation of mRNA synthesis at the TSS is an intricately regulated process that is influenced not only by the presence of VP30 but also by structure formation potential, length and sequence of the spacer region between PE1 and PE2. At the onset of this study, our state of knowledge was entirely based on the phenotype of the *NheI* HP mutant (Weik et al., 2002). The key conclusion of that study was that a destabilization of the NP HP structure particularly on the (+) RNA level - the result of 6 point mutations generating an *NheI* restriction site - makes mRNA synthesis VP30-independent, while the wt NP control is strictly VP30-dependent. This led the authors to suggest that stability of the NP hairpin structure defines the VP30 dependency of transcription (Weik et al., 2002). Using an enlarged series of mutant MGs and not only reporter assays but also qRT-PCR for sensitive quantifications of viral RNA species, we could extract several principles.

Regarding the *NheI* NP mutant, we were able to refine its phenotype. This mutant relaxes the VP30 dependency of transcription only partially (Fig. 5D, ~ 4-fold increased relative to the wt NP MG), but even the wt NP MG gives rise to residual VP30-independent transcription above background (Fig. 5D). Also, in our setup the mRNA levels were slightly reduced for the *NheI* NP mutant (+ VP30, Fig. 5D). These features likely escaped notice in the previous study (Weik et al., 2002) using Northern blot analysis of replicative MG RNA and MG reporter assays as experimental methods.

Most deviations from the NP HP structure, provided they did not further stabilize the stem structure, partially relaxed the VP30 dependency of transcription to different extents, but none made this process entirely VP30-independent as suggested previously (Weik et al., 2002). Even the  $\Delta 5'$  spacer mutant, lacking any potential to form secondary structures at the TSS on the genomic or antigenomic level, remained largely VP30-dependent (Fig. 3). These findings demonstrate that VP30 dependency of transcription can be relaxed to different extents upon sequence/structural changes of the native NP HP, but in all tested cases this relaxation remained partial. This may suggest that VP30 supports a conformational change in the L/VP35 polymerase complex that commits the enzyme to the transcription mode. With some variants of the PE1-PE2 spacer regions, such as construct VP35+1 (loop) v1, a larger fraction of polymerases



may be able to traverse this switch without VP30, but the conformational equilibrium cannot be entirely driven to the transcription-competent state without VP30. The transcription support function of VP30 requires the protein's capability to bind RNA, depends on an RNA-mediated interaction with VP35 and requires the dsRNA binding capacity of VP35 (Biedenkopf et al., 2016b, 2016a; Schlereth et al., 2016). VP30 was shown *in vitro* to bind RNA quite promiscuously, with a preference for ssRNA of mixed base composition in conjunction with an RNA secondary structure such as a hairpin (Schlereth et al., 2016). Promiscuous RNA binding of VP30 is consistent with our finding that VP30 was able to act on a variety of spacer variants that differed in sequence, structure and length. Moderate differences in VP30 affinity for the various 3'-leader RNAs have likely been masked by presumably saturating VP30 expression levels in the MG system.

Stabilization of the RNA secondary structure at the TSS, as in mutant NP G<sub>-72</sub> or VP40<sub>5</sub>, tightened the VP30 dependency, but simultaneously reduced overall transcription efficiency ( $\pm$  VP30) and also reduced cRNA synthesis (Fig. 4). A straightforward explanation is that fewer polymerases are able to productively bind the genomic promoter, possibly because VP35 has difficulties to resolve such stabilized duplex regions via its recently described RNA helicase activity (Shu et al., 2019). For variants NP G<sub>-72</sub> and VP40<sub>5</sub> (Fig. 4), VP30 dependency of transcription was maintained, but the increase of cRNA synthesis upon omission of VP30, as seen for the native 3'-leader, was attenuated. This suggests that structural stabilizations at the TSS prevent the polymerase complex from redirecting itself to the genome 3'-end for initiation of cRNA synthesis. Altogether, it has become evident from our findings that the NP hairpin structure at the TSS, which is slightly less stable than our NP G<sub>-72</sub> mutant, is optimized for viral RNA synthesis in terms of balancing transcription efficiency and tightness of VP30 dependency, as well as supporting the polymerase's capacity to increase cRNA synthesis when VP30 is absent.

Making the UN<sub>5</sub> hexamer phasing continuous between PE1 and PE2 (variants NP U<sub>-75</sub> and NP U<sub>-75</sub>/G<sub>-72</sub>) boosted transcription activity ( $\pm$  VP30). VP30-independent transcription reached the highest levels for these two mutants when normalized to the wt NP MG, as inferred from reporter activities in the RC MG and RD MG background as well as mRNA levels determined by qRT-PCR (Fig. 5B-D). Furthermore, cRNA synthesis was increased in the presence of VP30, but the switch from transcription to replication in the absence of VP30 was dysregulated (Fig. 5; Bach et al., 2019a). The phenotype of these two mutants emphasizes the important function of UN<sub>5</sub> hexamer phasing in

supporting RdRp activity at the 3'-leader promoter, which is likely intertwined with NP:RNA binding (Bach et al., 2019a). The very similar phenotype of mutants NP U<sub>-75</sub> and NP U<sub>-75</sub>/G<sub>-72</sub> is surprising, considering that the additional G<sub>-72</sub> mutation largely compensated the loss of structure caused by the U<sub>-75</sub> mutation (Fig. 5A).

A closer look at the predicted structures of the wt NP HP in comparison with the NheI NP, NP U<sub>-75</sub> and NP U<sub>-75</sub>/G<sub>-72</sub> variants suggests the possibility that sequence and local stabilities of the lower stem region might play a role in VP30 dependency (Fig. 5A). All three mutant variants are predicted to include nucleotides of PE1 and PE2 in their basal stem region on the genomic RNA level. Provided these structures indeed form (transiently), this might perturb recognition of signals involved in the control of transcription or replication. Since the phenotype of the NheI NP mutant was less pronounced than those of NP U<sub>-75</sub> and NP U<sub>-75</sub>/G<sub>-72</sub>, other features must play a role as well. Nevertheless, changes in the position of the lower HP stem relative to PE1 and PE2 are also predicted for the L+1 (stem) and VP35+1 (stem) variants, and transcriptional VP30 dependency was indeed tightened for these two constructs relative to their corresponding loop variants (Fig. 3). However, this correlation needs to be further explored, as the VP30+4 (stem) variant with a 4-nt insertion between the stem and PE2 did not show such stringent regulation of transcription by VP30 (Fig. 2 and 3).

In variants NP U<sub>-75</sub>, NP U<sub>-75</sub>/G<sub>-72</sub>, VP35+1 (stem) and Δ5' spacer, UN<sub>5</sub> hexamer phasing is continuous, but VP30 relaxation of transcription was weak for the VP35+1 (stem), moderate for the Δ5' spacer and high for the U<sub>-75</sub> mutants. This illustrates that UN<sub>5</sub> hexamer periodicity between PE1 and PE2 per se is insufficient to explain the strong phenotype of mutants NP U<sub>-75</sub> and U<sub>-75</sub>/G<sub>-72</sub>. Further investigations are required to understand the UN<sub>5</sub> hexamer code, its sequence constraints and the molecular functions it encodes.

Low residual mRNA, cRNA and vRNA synthesis can occur in the case of MG constructs violating the rule of 6 [variants VP35, NP-1 (stem)]. The extent depends on the sequence/structure context and low-efficiency mRNA synthesis is stimulated by VP30 (Fig. 6). It is unclear why the VP35 construct was more efficient than the NP-1 (stem) construct. One possibility is that the 1-nt deletion in the latter construct changes the HP stem distance to PE2 as discussed above. Other mutants, such as NP mutants with the 1-nt deletion in the loop, could be analyzed to test this possibility. Another testable possibility is that the 3<sup>rd</sup> base pair of the genomic HP stem, which is C:G in the NP and A:U in the VP35 HP (Fig. 6A), plays a key role in this context.

Data obtained with the GP-2 ( $\Delta 2$  nt PE1) mutant MG suggest that elimination of genome nucleotides -54 and -55 and thereby shortening of the distance between the genome 3'-end and TSS prevented any elevation of mRNA synthesis by VP30 to levels above background (Fig. 6B). This would be consistent with the notion that (i) nt -54 and -55 are part of the EBOV transcription promoter and necessary for binding of the transcriptase complex, including VP30, (ii) that a certain spacing to the genome 3'-end is key to viral transcription, or (iii) that the first nucleotide (C<sub>-56</sub>) needs to be in a certain NP phase as was suggested for the paramyxovirus SeV (Bach et al., 2019c; Cordey and Roux, 2007).

Expansions of up to 66 nt in the spacer region still allow transcription, although at reduced levels and not specifically affecting VP30-dependent or VP30-independent activity. This suggests that the fraction of RdRp complexes that productively interact with the 3'-leader promoter decreases when the spacing between PE1 and PE2 exceeds a certain length limit (Bach et al., 2019a). Yet, at those RdRp complexes that succeed to initiate transcription, VP30 can exert its transcriptional support function.

## CONCLUSIONS

Our study demonstrates that any changes in the PE1-PE2 spacer of the EBOV 3'-leader promoter in terms of RNA sequence, structure formation potential and length alter either the tightness of VP30 control of transcription, and/or overall efficiency of transcription (and replication), and/or dampen the induction of replication in the absence of VP30. Thus, the native spacer region is evolutionary optimized for balanced and efficient viral transcription and replication. The molecular basis, yet far from being comprehensively understood, is an intricate interplay of RNA sequence signals, RNA structure, spacer length and UN<sub>5</sub> hexamer periodicity in the 3'-leader promoter.

## MATERIAL AND METHODS

Cell culture, EBOV-specific minigenome assays, Renilla and Firefly luciferase activity assays as well as total RNA extraction from cells transfected with plasmids of the EBOV minigenome system and 2-step strand-specific qRT-PCR of EBOV MG RNAs were carried out exactly as described (Bach et al., 2019a).

### *Plasmids*

Plasmids encoding the wt replication-competent (RC) and replication-deficient (RD) MG, mutant variants thereof, as well as pCAGGS-derived plasmids coding for the Zaire EBOV nucleocapsid proteins (VP30, NP, VP35 and L) and the T7 RNA polymerase were described earlier (Hoenen et al., 2006, 2010; Bach et al., 2019a, 2019b). A Firefly luciferase-encoding pGL4.13 plasmid, used as transfection control, was purchased from Promega.

### *Quantification and statistical analysis*

All statistical analyses were performed using GraphPad Prism version 8.1.1 for Windows. Samples were analyzed by applying the unpaired *t* test with Welch's correction and a statistical significance level of 0.05. Calculated *p* values and definitions are indicated in the respective figure legends.

## FUNDING

This work was supported by the German Research Foundation (DFG), grant CRC 1021, to S. Becker and R. K. Hartmann.

## ACKNOWLEDGEMENTS

We like to acknowledge technical support by Astrid Herwig (cell culture) and Arnold Grünweller for fruitful discussions.

## REFERENCES

- Albariño, C. G., Guerrero, L. W., Chakrabarti, A. K., and Nichol, S. T. (2018). Transcriptional analysis of viral mRNAs reveals common transcription patterns in cells infected by five different filoviruses. *PLoS One* 13, 1–13. doi:10.1371/journal.pone.0201827.
- Biedenkopf, N., Hartlieb, B., Hoenen, T., and Becker, S. (2013). Phosphorylation of Ebola Virus VP30 Influences the Composition of the Viral Nucleocapsid Complex. *J. Biol. Chem.* 288, 11165–11174. doi:10.1074/jbc.m113.461285.
- Biedenkopf, N., Lier, C., and Becker, S. (2016a). Dynamic Phosphorylation of VP30 Is Essential for Ebola Virus Life Cycle. *J. Virol.* 90, 4914–4925. doi:10.1128/jvi.03257-15.
- Biedenkopf, N., Schlereth, J., Grünweller, A., Becker, S., and Hartmann, R. K. (2016b). RNA Binding of Ebola Virus VP30 Is Essential for Activating Viral Transcription. *J. Virol.* 90, 7481–7496. doi:10.1128/jvi.00271-16.
- Calain, P., Monroe, M. C., and Nichol, S. T. (1999). Ebola Virus Defective Interfering Particles and Persistent Infection. *Virology* 262, 114–128. doi:10.1006/VIRO.1999.9915.
- Cordey, S., and Roux, L. (2007). Further characterization of a paramyxovirus transcription initiation signal: Search for required nucleotides upstream and importance of the N phase context. *J. Gen. Virol.* 88, 1555–1564. doi:10.1099/vir.0.82701-0.
- Deflubé, L. R., Cressey, T. N., Hume, A. J., Olejnik, J., Haddock, E., and Feldmann, F. (2019). Ebolavirus polymerase uses an unconventional genome replication mechanism. 116, 8535–8543. doi:10.1073/pnas.1815745116.
- Feldmann, H., Mühlberger, E., Randolph, A., Will, C., Kiley, M. P., Sanchez, A., et al. (1992). Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus Res.* 24, 1–19. doi:10.1016/0168-1702(92)90027-7.
- Hoenen, T., Groseth, A., Kolesnikova, L., Theriault, S., Ebihara, H., Hartlieb, B., et al. (2006). Infection of Naïve Target Cells with Virus-Like Particles: Implications for the Function of Ebola Virus VP24. *J. Virol.* 80, 7260–7264. doi:10.1128/JVI.00051-06.
- Hoenen, T., Jung, S., Herwig, A., Groseth, A., and Becker, S. (2010). Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription. *Virology* 403, 56–66. doi:10.1016/j.virol.2010.04.002.
- Ivanov, A., Ramanathan, P., Parry, C., Ilinykh, P. A., Lin, X., Petukhov, M., et al. (2019). Global phosphoproteomic analysis of Ebola virions reveals a novel role for VP35 phosphorylation-dependent regulation of genome transcription. *Cell. Mol. Life Sci.*, 1–25. doi:10.1007/s00018-019-03303-1.
- Mühlberger, E. (2007). Filovirus replication and transcription. *Future Virol.* 2, 205–215. doi:10.2217/17460794.2.2.205.
- Mühlberger, E., Trommer, S., Funke, C., Volchkov, V., Klenk, H.-D., and Becker, S. (1996). Termini of all mRNA species of Marburg virus: Sequence and secondary

- structure. *Virology* 223, 376–380. Available at: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed4&NEWS=N&AN=1996293636>.
- Mühlberger, E., Weik, M., Volchkov, V. E., Klenk, H. D., and Becker, S. (1999). Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J. Virol.* 73, 2333–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9971816><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC104478>.
- Rougeron, V., Feldmann, H., Grard, G., Becker, S., and Leroy, E. M. (2015). Ebola and Marburg haemorrhagic fever. *J. Clin. Virol.* 64, 111–119. doi:10.1016/j.jcv.2015.01.014.
- Sanchez, A., and Kiley, M. P. (1987). Identification and analysis of ebola virus messenger RNA. *Virology* 157, 414–420. doi:10.1016/0042-6822(87)90283-2.
- Sanchez, A., Kiley, M. P., Holloway, B. P., and Auperin, D. D. (1993). Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res.* 29, 215–240. doi:10.1016/0168-1702(93)90063-S.
- Schlereth, J., Grünweller, A., Biedenkopf, N., Becker, S., and Hartmann, R. K. (2016). RNA binding specificity of Ebola virus transcription factor VP30. *RNA Biol.* 13, 783–798. doi:10.1080/15476286.2016.1194160.
- Shabman, R. S., Jabado, O. J., Mire, C. E., Stockwell, T. B., Edwards, M., Mahajan, M., et al. (2014). Deep Sequencing Identifies Noncanonical Editing of Ebola and Marburg Virus RNAs in Infected Cells. 5, 1–11. doi:10.1128/mBio.02011-14.Editor.
- Shu, T., Gan, T., Bai, P., Wang, X., Qian, Q., Zhou, H., et al. (2019). Ebola virus VP35 has novel NTPase and helicase-like activities. *Nucleic Acids Res.* 47, 5837–5851. doi:10.1093/nar/gkz340.
- Weik, M., Enterlein, S., Schlenz, K., and Muhlberger, E. (2005). The Ebola Virus Genomic Replication Promoter Is Bipartite and Follows the Rule of Six. *J. Virol.* 79, 10660–10671. doi:10.1128/jvi.79.16.10660-10671.2005.
- Weik, M., Modrof, J., Klenk, H.-D., Becker, S., and Mühlberger, E. (2002). Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J. Virol.* 76, 8532–9. doi:10.1128/JVI.76.17.8532.
- Xu, W., Luthra, P., Wu, C., Batra, J., Leung, D. W., Basler, C. F., et al. (2017). Ebola virus VP30 and nucleoprotein interactions modulate viral RNA synthesis. *Nat. Commun.* 8, 1–11. doi:10.1038/ncomms15576.

## FIGURE LEGENDS

**Fig. 1:** (A) Architecture of the negative sense EBOV genome (vRNA). Potential secondary structures at the transcription start site (TSS) and internal gene starts (GS) are depicted at the top. Corresponding hairpin structures are also predicted to form on the mRNA and antigenomic RNA level (not shown). The TSS of the first NP gene and GS signals of the six internal EBOV genes are highlighted in light blue. Schematic white boxes mark the reading frames for EBOV proteins NP, VP35, VP40, GP, VP30, VP24 and L; 5'- and 3'-UTRs are depicted as light gray boxes, with dark gray areas marking the position of the predicted secondary structures illustrated above the genome; the box representing the first NP hairpin structure is highlighted in red; mRNAs are illustrated as colored lines, with dots indicating the 5'-cap; the number of lines below each gene symbolize (without representing quantitative ratios) the decrease in the amount of mRNA transcripts from the first NP to the last L mRNA; p, genomic 3'-leader and antigenomic trailer promoters (Calain et al., 1999). (B) Illustration of the replication-competent (RC; scheme at the top) and replication-deficient (RD; bottom scheme) minigenome (MG) used in this study; in the scheme at the top, the secondary structure potential at the genomic 3'-end is depicted, which was validated by structure probing of naked RNA *in vitro* (Weik et al., 2002 & 2005; Schlereth et al., 2016). RD MGs lack the trailer promoter's terminal 55 nt (red line), thus restricting RNA synthesis to transcription and antigenome synthesis. PE, promoter elements (in green); the genomic replication promoter is bipartite, consisting of PE1 and PE2. Bold pink letters highlight the 3'-U residues of 3'-UN<sub>5</sub> hexamers in PE1 and PE2 shown to be crucial for viral replication (Weik et al., 2005) and transcription (Bach et al., 2019a). The NP GS signal (TSS) and PE2 are separated by a spacer region (orange letters). Negative values are used for nucleotide numbering of the genomic (-) RNA, reflecting the genome's 3' to 5' negative-sense orientation. The genomic 3'-terminal G is depicted in parentheses as its presence is not essential for initiation of antigenomic RNA or leader transcript synthesis by the EBOV RNA polymerase complex (Deflubé et al., 2019; Bach et al., 2019c).

**Fig. 2:** Schematic illustration of MG variants analyzed in Fig. 3. The native (wt) NP hairpin (HP) structure predicted on the genomic RNA level is shown in the center (area shaded in light blue); the NP HP was either changed by mutation (variant NheI NP) or deletion (variant  $\Delta 5'$  spacer), or was replaced with (engineered) HP structures derived from GS regions of internal EBOV genes. For color coding, see legend to Fig. 1B. Hexamer periodicity between PE1 and PE2 is depicted for each variant by highlighting the pink 3'-U residues of 3'-UN<sub>5</sub> hexamers (Bach et al., 2019a). MG variants with HP structures derived from GS regions of internal EBOV genes (except for the VP40 HP)

were adjusted to hexamer phasing by inserting (boxed nt) or deleting nucleotides, either at the 5'-side of the hairpin stem or in its apical loop; ins., insertion; del., deletion. Underlined nucleotides in the NheI NP hairpin mark the six point mutations (mut.) relative to the wt NP HP.

**Fig. 3:** Reporter gene activity of **(A)** RC MG or **(B)** or RD MG variants specified in Fig. 2 and transfected together with a plasmid expressing VP30 (white bars) or omitting such a plasmid (light blue bars). The upper dotted horizontal lines represents the respective activity value for the MG carrying the native 3'-leader (wt NP), which was set to 100 % (gray columns). The lower dotted horizontal line marks the normalized activity value (- VP30) for the respective MG harboring the destabilized NheI NP HP (Fig. 2). This mutant, which was previously reported to relax the VP30 requirement for EBOV transcription (Weik et al., 2002), was used as reference; -L background control: MG transfection without the plasmid encoding the polymerase L gene (dark gray columns). Values ( $\pm$  standard error of the mean, SEM) are based on at least 3 biological replicates with 2 or 3 technical replicates each. For the reference NheI NP mutant (RD MG), the mean *n*-fold change of VP30-independent activity compared to the wt NP RD MG is indicated in red in panel B (1.6-fold). **(C)** qRT-PCR analysis of mRNA levels in selected samples obtained from the same RC MG-transfected cells as analyzed in panel A. Mean  $2^{-\Delta\Delta CT}$  values as a measure of mRNA levels relative to the wt NP MG, based on at least 3 independent experiments with a total of at 2 or 3 technical replicates each. Red numbers indicate the mean *n*-fold change relative to the wt NP MG in the absence of VP30. For details on the qRT-PCR setup, see "Strand-specific qRT-PCR" in Material and Methods. \*\**p* < 0.01; \*\*\*\**p* < 0.0001 (unpaired *t* test with Welch's correction).

**Fig. 4:** **(A)** RNA secondary structure potential of stabilized variants of the wt NP HP (variant NP G<sub>-72</sub>) or the VP40 HP (variants VP40<sub>5</sub> and VP40<sub>6</sub>). Potential structures are shown on the genomic (vRNA) or mRNA level;  $\Delta G$  (kcal/mol) values of the centroid structures predicted by RNAfold using the default parameters (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>) are depicted in red color below (vRNA) or above (mRNA) the corresponding secondary structure. The genomic TSS signal is marked in light blue, point mutations (relative to the parental native structure) are marked in pink, nt insertions in green and deletions (del.) by arrows. **(B-E)** Influence of RNA secondary structure stabilization at the TSS on reporter activity and viral RNA synthesis in the presence and absence of VP30. **(B)** Luciferase reporter gene assays of the HP constructs illustrated in panel A; the column color code is explained at the bottom of the figure. Activities of the wt NP MG were set to 100 % in each experiment. Mean relative reporter gene activity values  $\pm$  SEM are derived from at



least 3 independent experiments with 3 technical replicates each. **(C-E)** Corresponding qRT-PCR analyses for the quantification of mRNA, cRNA and vRNA levels changes in total RNA samples derived from the same MG-transfected cells as analyzed for reporter activity in panel B. Presented values are based on 3 independent experiments with 3 technical replicates each.  $2^{-\Delta\Delta CT}$  values were normalized to RNA levels measured for the wt NP MG ( $2^{-\Delta\Delta CT} = 1$ ). \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$  (unpaired  $t$  test with Welch's correction).

**Fig. 5:** Effects of NP hairpin destabilization or elimination as well as continuous hexamer periodicity between PE1 and PE2 on reporter activity and viral RNA levels in the presence versus absence of VP30. **(A)** Predicted secondary structures on the genomic and mRNA level for the wt NP HP (bottom), the NheI NP HP and  $\Delta 5'$  spacer variants (left box at the top) as well as variants NP U<sub>-75</sub> and NP U<sub>-75</sub>/G<sub>-72</sub> (right box at the top). For details on color coding, see Fig. 1B and Fig. 2.  $\Delta G$  (kcal/mol) values of the centroid structures predicted by RNAfold using the default parameters are indicated. Mutated nucleotide positions are underlined;  $\Delta$ , deletion. Horizontal arrows additionally highlight the point mutation at pos. -75/+75 (red) and the stabilizing point mutations at -72/+72 (black). For nucleotide numbering, negative values [genomic (-) RNA] or positive values [mRNA/ (+) RNA] were used, starting at the genome 3'-end. **(B, C)** Luciferase reporter gene assay of RC **(B)** or RD **(C)** MG variants described in Fig. 5A. Reporter activities relative to those of the wt NP MG (set to 100 %) are given, based on at least 3 biological replicates ( $\pm$  SEM) with 2 or 3 technical replicates each. Red numbers in panel C represent mean  $n$ -fold changes of activity relative to the wt NP MG in the absence of VP30. **(D-F)** Corresponding qRT-PCR analyses of total RNA samples derived from the same cells as analyzed in panel B, specific for **(D)** mRNA, **(E)** antigenomic/copy RNA (cRNA) and **(F)** genomic/viral RNA (vRNA). Mean  $2^{-\Delta\Delta CT}$  values  $\pm$  SEM are derived from at least three independent experiments with two or three technical replicates each. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$  (unpaired  $t$  test with Welch's correction).

**Fig. 6:** Effects of VP30 on MGs violating the rule of 6 between PE1 and PE2. **(A)** Predicted TSS-spacer structures of MG variants that either conform to [wt (NP); VP35+1 (loop) v1] or violate [VP35; NP-1 (stem); GP-2 ( $\Delta 2$  nt PE1)] the rule of six. In variant GP-2 ( $\Delta 2$  nt PE1), genome nucleotides -54 and -55 of PE1 are deleted (Bach et al., 2019a). Insertions or deletions of nucleotides are marked by black boxes. ins, insertion; del, deletion. For details on color code, see legend to Fig. 1B. **(B-D)** Corresponding qRT-PCR analyses of total RNA obtained from cells transfected with the RC MGs illustrated in panel A, distinguishing between relative amounts of

(B) viral mRNA, (C) cRNA and (D) vRNA. Mean  $2^{-\Delta\Delta CT}$  values  $\pm$  SEM are derived from at least 3 independent experiments with 2 or 3 technical replicates each. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\*\* $p < 0.0001$  (unpaired  $t$  test with Welch's correction).

**Fig. 7:** Influence of hairpin/spacer length on VP30-independent transcription. (A) Luciferase reporter gene assays of monocistronic RC MGs comprising the VP40 HP and mutant variants thereof. The VP40 hairpin was elongated in steps of 6 nt. MG variants VP24-4, L+1 (stem) and L+1 (loop) were included as well to represent different sequence contexts. Numbers above the columns indicate the number of nucleotides that were inserted into the spacer relative to the wt NP MG. Numbers are shown in pink for variants with a different sequence context but the same spacer expansion of 48 nt. (B) Illustration of the hairpin structures predicted to form at the mRNA 5'-ends of the MGs analyzed in panel A.  $\Delta G$  values are those of the centroid structure predicted by RNAfold using the default parameters. Nucleotide inserted into the EBOV-specific hairpin sequences are highlighted in green. Nucleotide positions in hairpin stems that deviate from the wt NP stem sequence are marked in red. Lower stem regions are highlighted by the light yellow boxes for comparison with the NP HP stem.



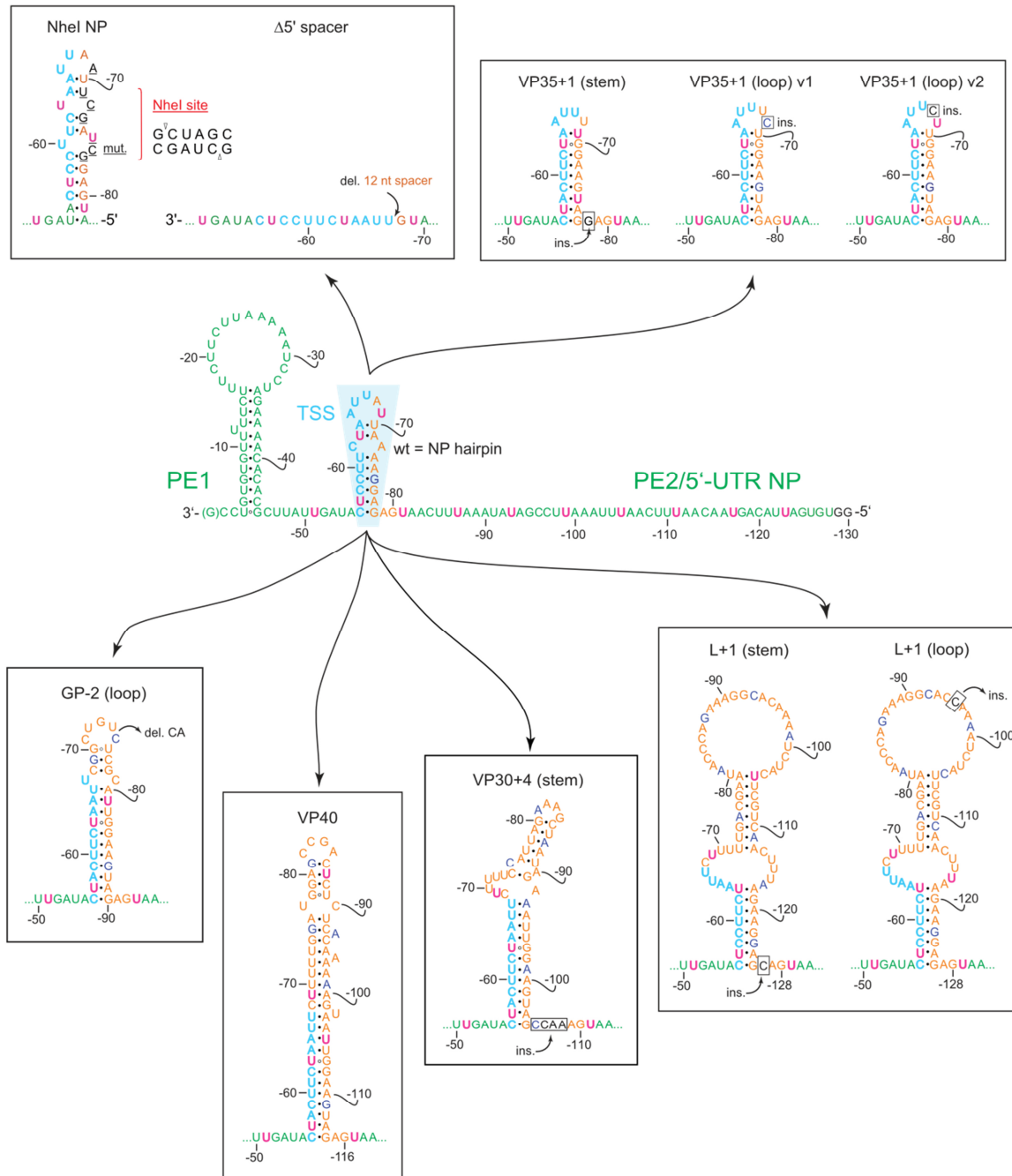
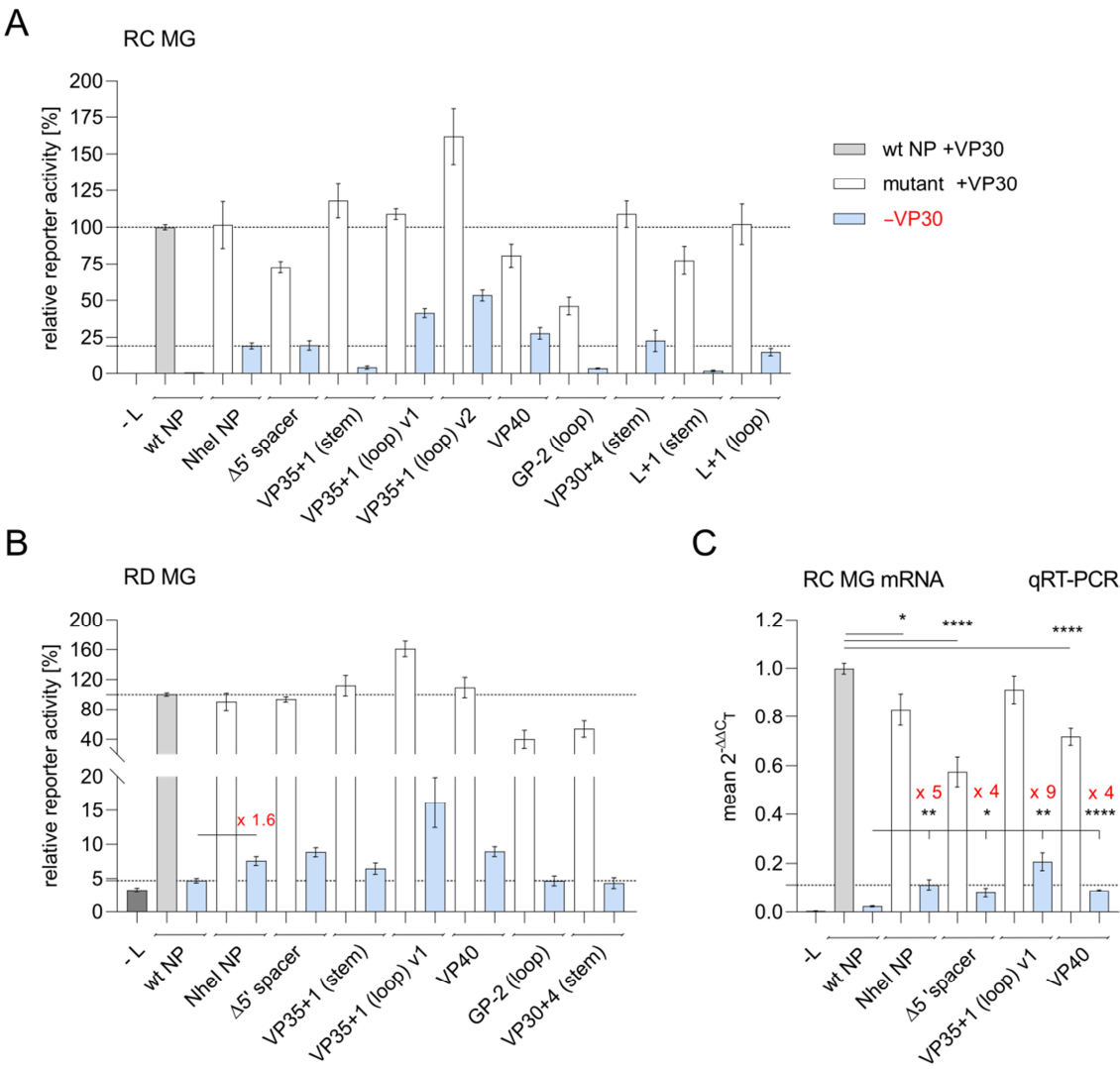


Fig. 2



A

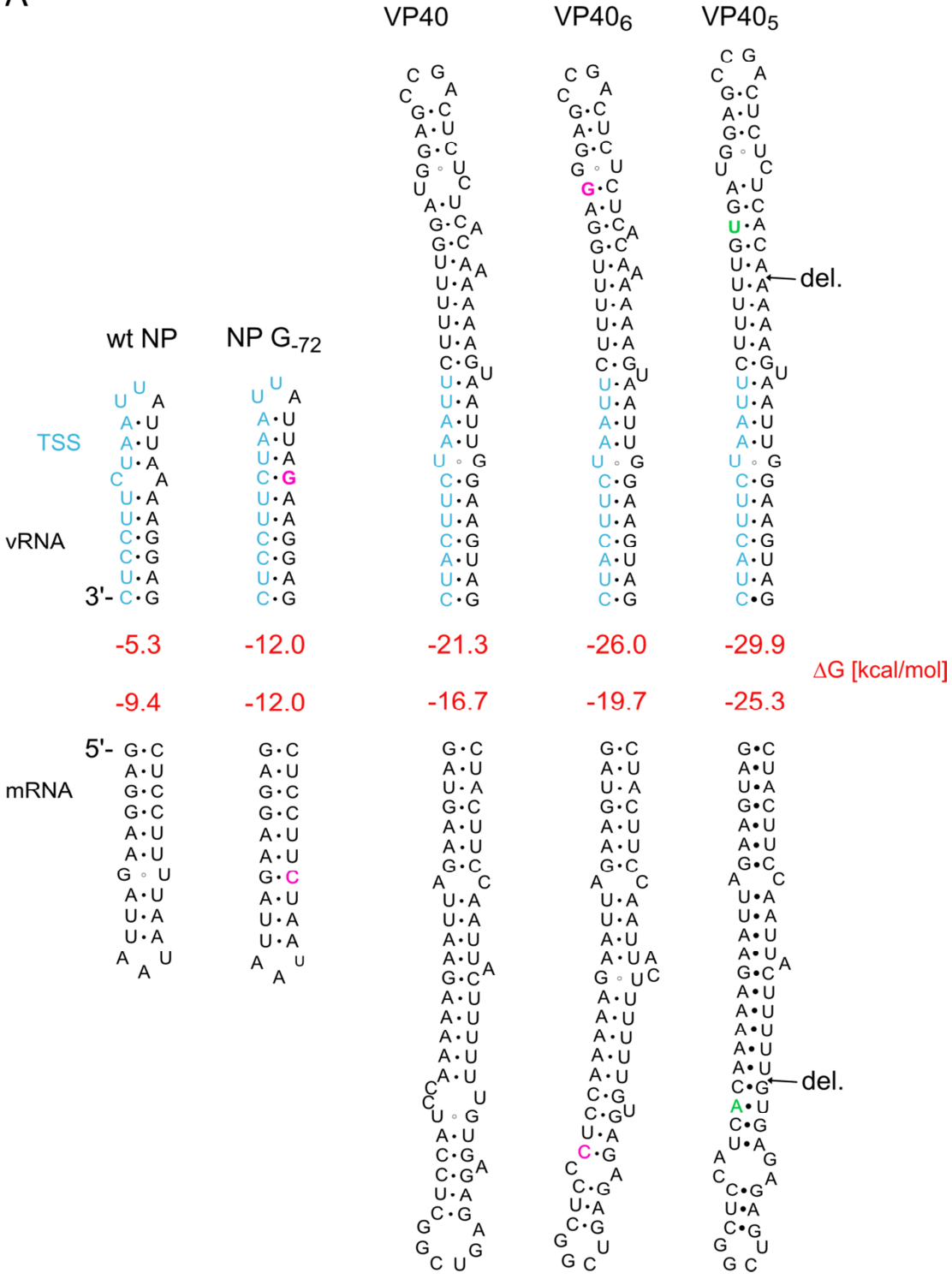


Fig. 4 A

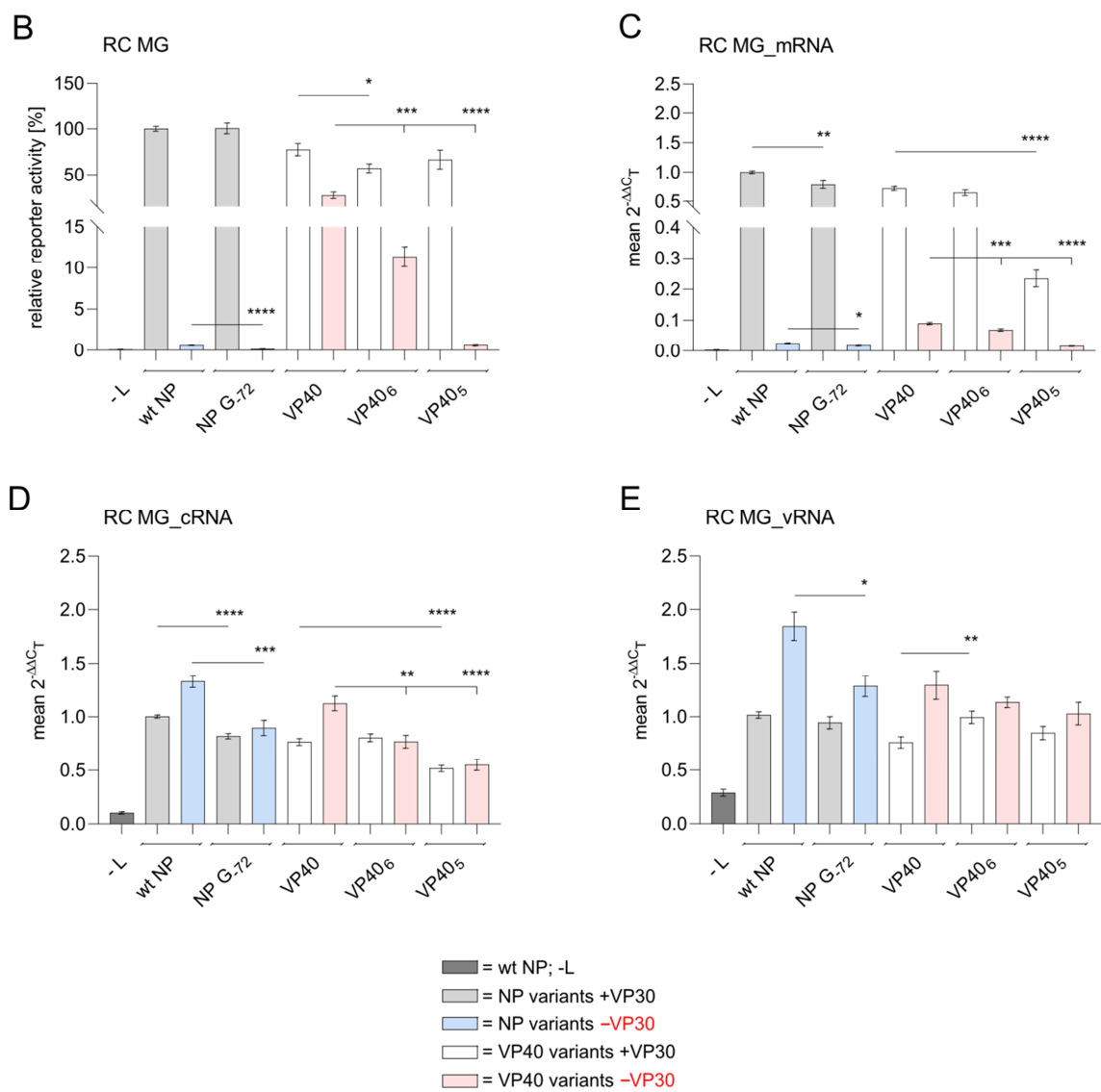


Fig. 4 B-E

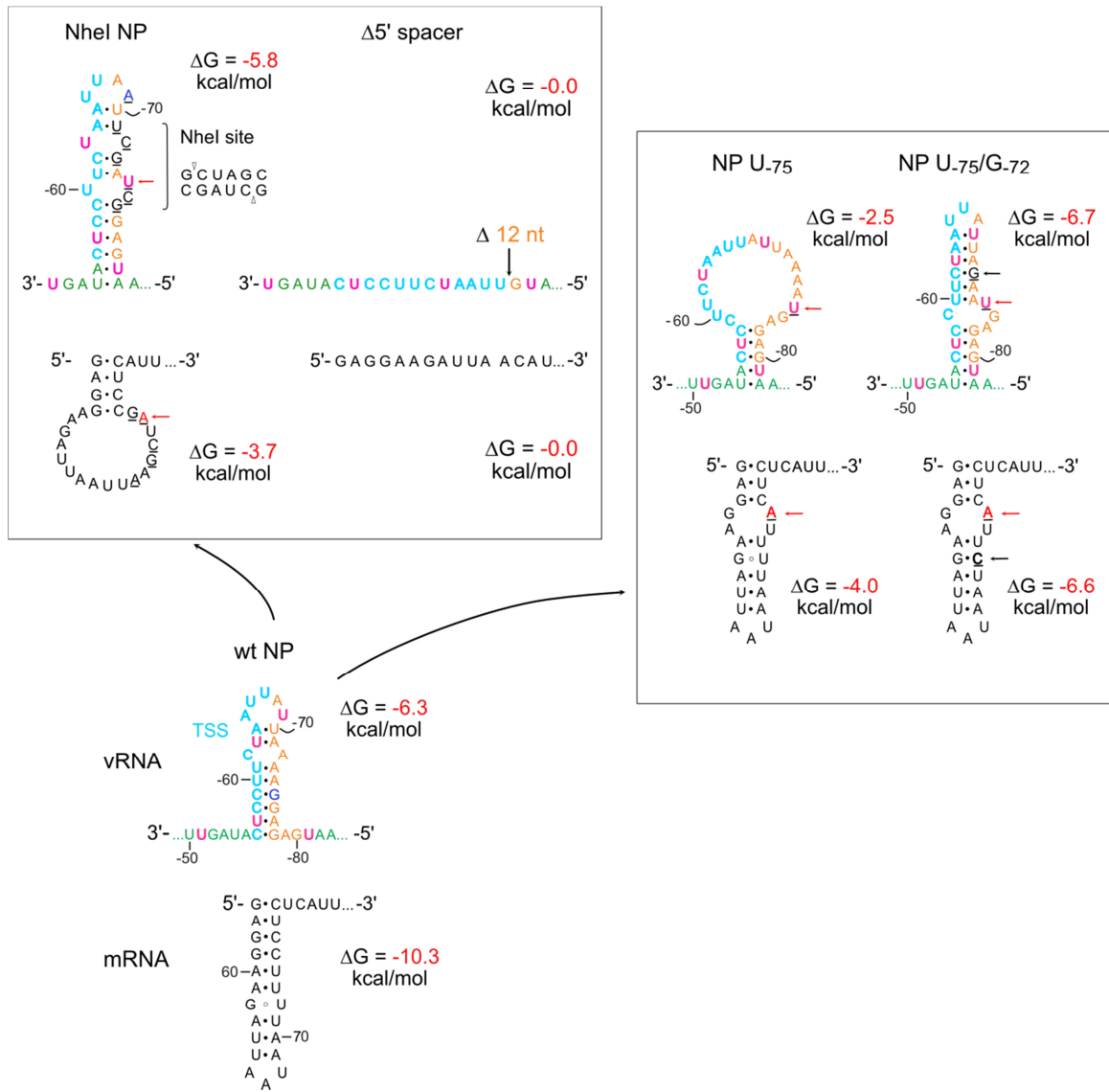


Fig. 5 A



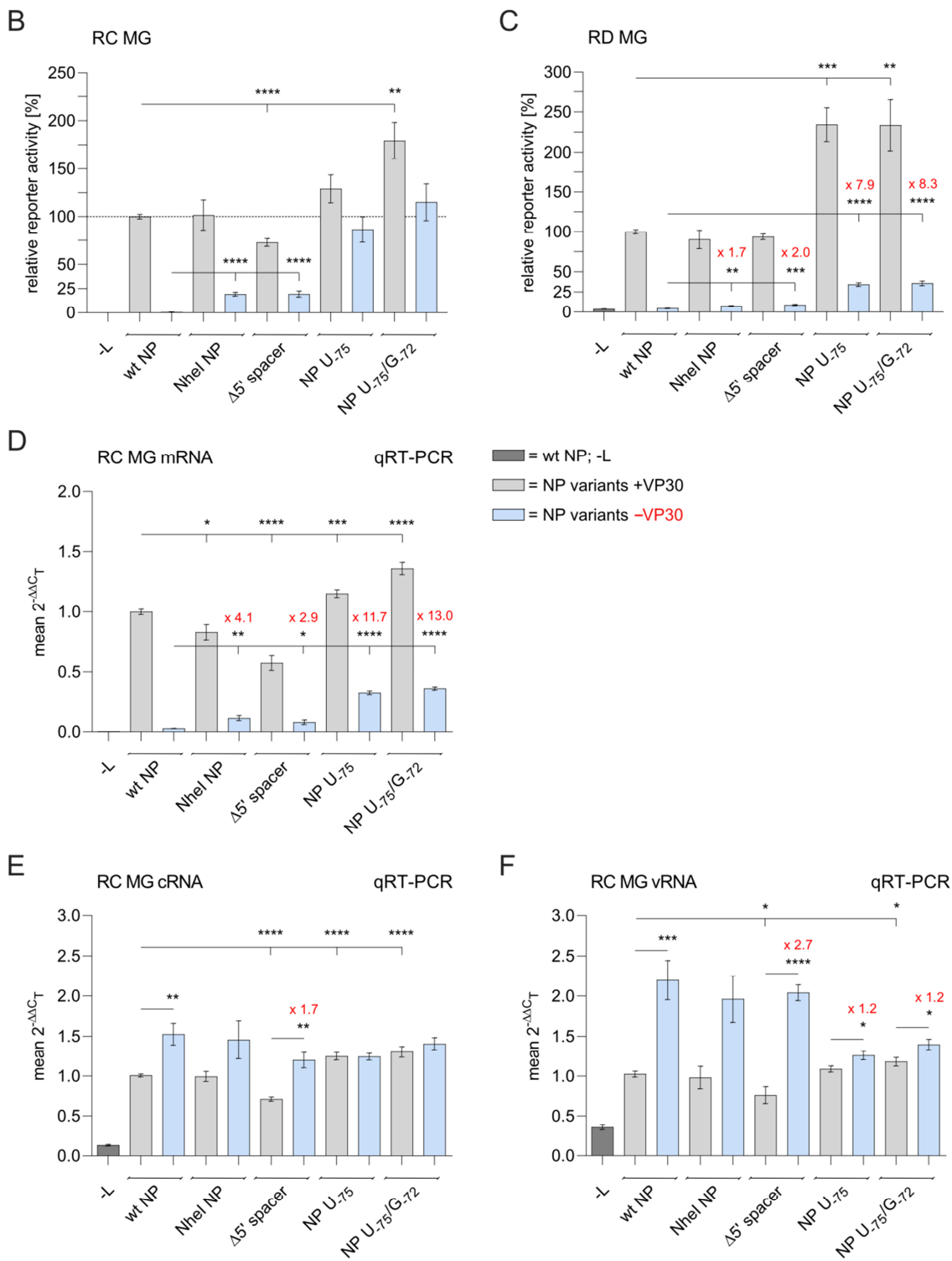


Fig. 5 B-F

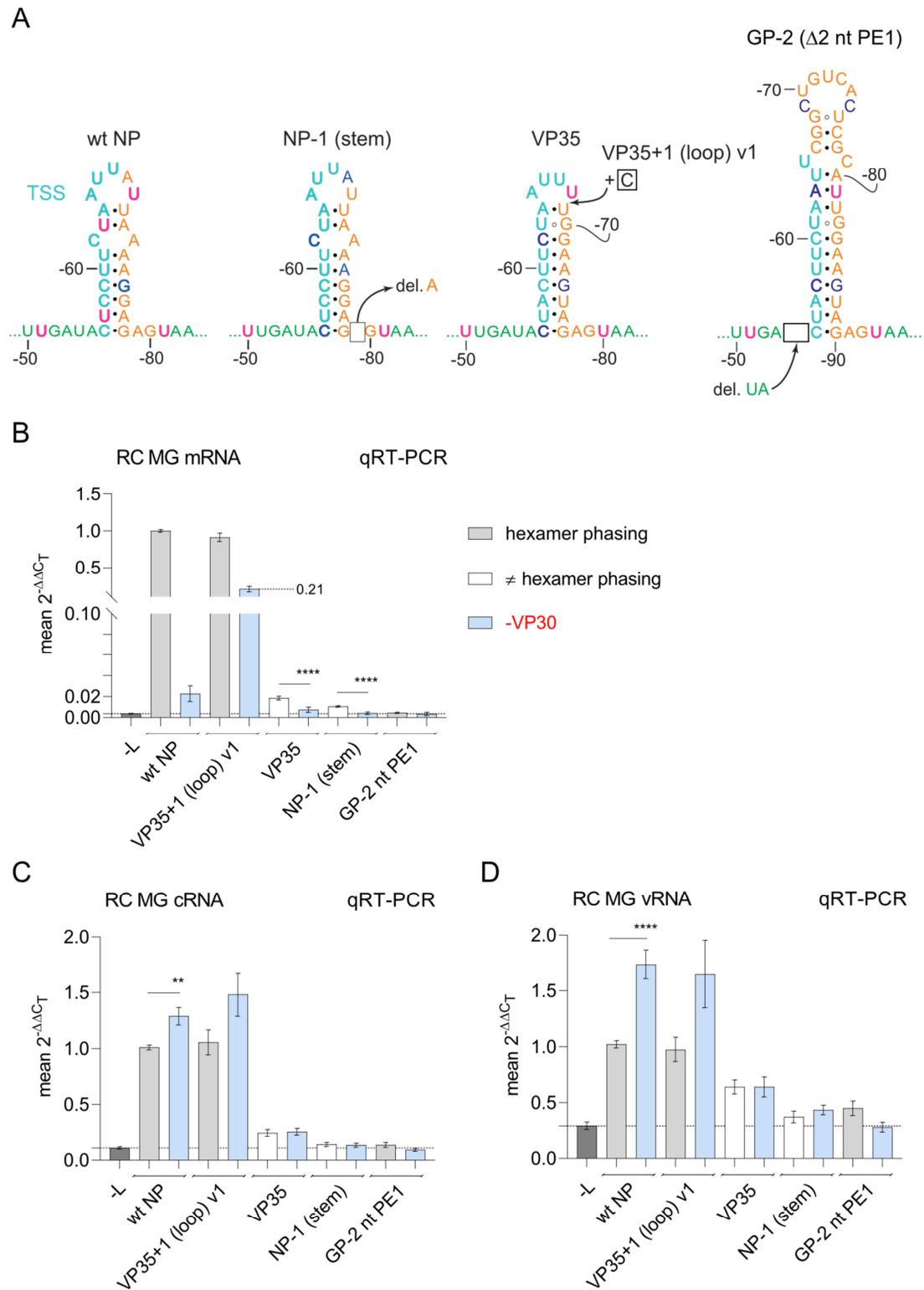


Fig. 6

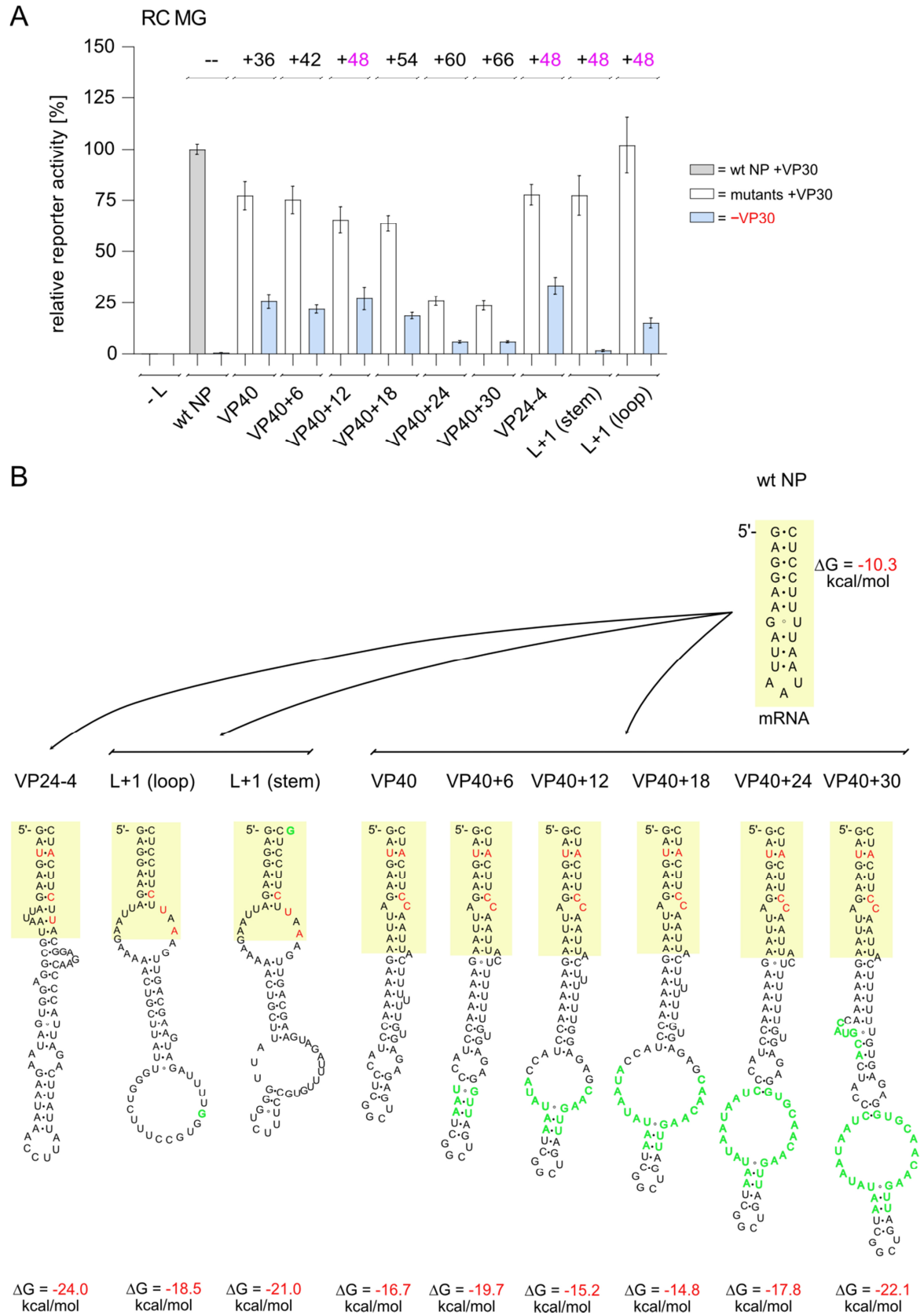
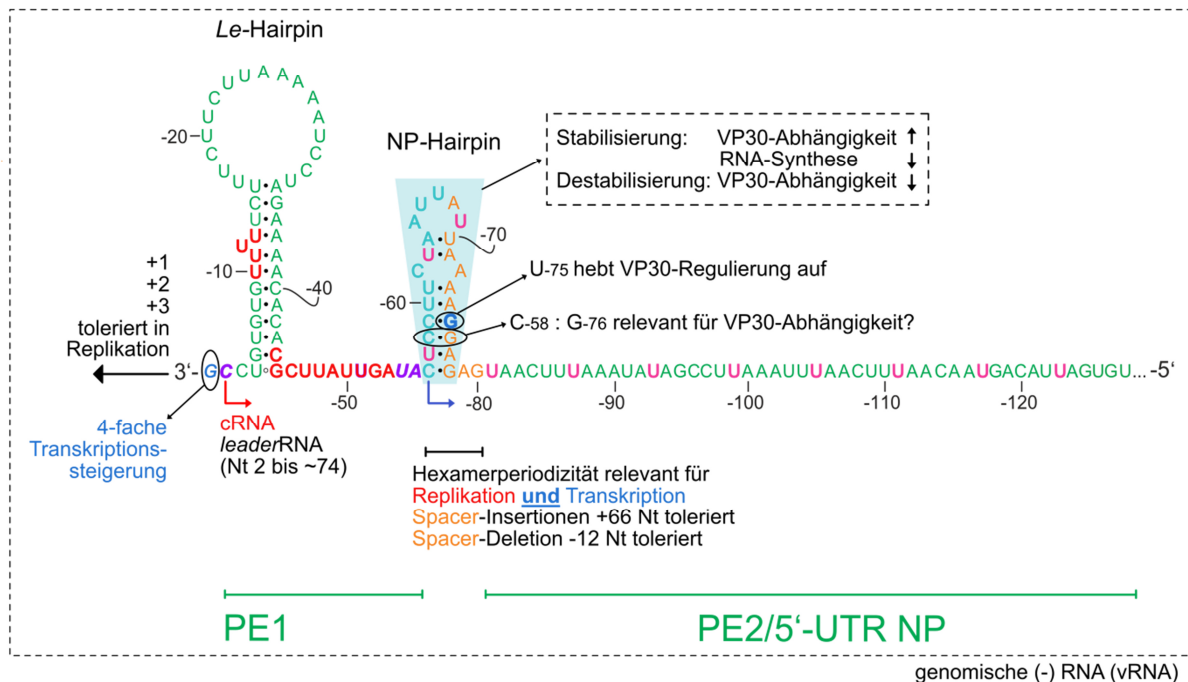


Fig. 7

## 5 DISKUSSION UND ZUSAMMENFASSUNG



**Abb. 18: Zusammenfassung regulatorischer Sequenz- und Strukturbereiche des genomischen 3'-Leader-Promotors.** Das Farbschema ist grundsätzlich das Gleiche wie in Abb. 10. Rot markierte, fettgedruckte Nukleotide (Nt) sind essenziell für die Replikation. Violette, kursivgedruckte Nt (C-2, U-54, A-55) sind gleichermaßen für Replikation und Transkription notwendig. G-1 (blau, kursiv) ist transkriptionssteigernd (~ 4-fach), jedoch nicht obligat für die virale Replikation. Die cRNA- und leaderRNA-Synthesen werden exklusiv an C-2 initiiert (roter Pfeil). Die NP mRNA-Synthese erfolgt spezifisch an C-56 (blauer Pfeil). Die Replikationspromotorelemente PE1 und 2 werden durch das erste (NP)-Transkriptionsstartsignal (Nt -56 bis -67; cyan) und einen Spacer-Bereich (Nt -68 bis -80; orange) separiert. TSS und Spacer bilden eine Haarnadelstruktur (NP-Hairpin) aus, die die VP30-Abhängigkeit der Transkription reguliert. Die Nukleotide U-51 (blaue Umrahmung), U-57, U-63 und U-69 begründen ein vermutetes transkriptionsrelevantes 3'-UN<sub>5</sub>-Hexamerraster. Dieses Raster wird durch das G-75 vom replikationsrelevanten UN<sub>5</sub>-Raster in PE2 (Nt -81 bis -128) getrennt. Ein kontinuierliches UN<sub>5</sub>-Raster von U-51 - U-128 durch Einführung einer G-75 zu U-Punktmutation steigert (weitgehend) sekundärstrukturunabhängig die allgemeine virale RNA-Synthese - jedoch auf Kosten der Feinregulation von Transkription und Replikation. Die Initiation der cRNA-, leaderRNA- und mRNA-Synthesen erfolgt nur bei Einhaltung der Spacer-„Rule of 6“ (Hexamerperiodizität), d.h. ausschließlich Insertionen und Deletionen eines Vielfachen der Zahl 6 werden von der viralen Polymerase toleriert.

Vom 3'-Leader-Promotor werden sowohl die virale Antigenom (cRNA)-Synthese als auch die Transkription der ersten (NP) mRNA initiiert. Somit liegt eine stete Konkurrenzsituation der viralen Transkription und Replikation am genomischen 3'-Promotor vor. Besonders relevant ist dies für die späte Infektionsphase, in der Replikation und sekundäre Transkription parallel stattfinden. Für eine optimale Virusvermehrung ist es essenziell, dass beide Prozesse ausgewogen aufeinander abgestimmt werden. Die Existenz eines mit PE1 überlappenden Transkriptionspromotors war vor Beginn der vorliegenden Studie rein spekulativ. Im Rahmen dieser Studie konnten einige sequenzspezifische und strukturelle Voraussetzungen der Polymerisationsinitiation am genomischen 3'-Leader-Promotor sowie mechanistische Aspekte der Regulation von Transkription und Replikation durch das virale Protein 30 (VP30) geklärt werden (Abb. 18).

Die vorliegenden Ergebnisse lassen die Hypothese zu, dass der EBOV-Transkriptionspromotor (fast) den gesamten PE1-Bereich umfasst. Dies lässt sich aus Deletionsanalysen des Replikations-defizienten Minigenoms ableiten, die die Nukleotide C<sub>-2</sub> sowie U<sub>-54</sub> und A<sub>-55</sub> betrafen. Deletionen dieser Nukleotide eliminierten die virale Transkription vollständig. Das erste Genomnukleotid (G<sub>-1</sub>) ist für die Replikation verzichtbar, spielt aber eine deutliche Rolle für die virale Transkription, da seine Anwesenheit die Transkription ca. 4-fach steigert. Eine vorausgehende Studie von L. Deflubé und Koautoren zeigte, dass die EBOV-cRNA-Synthese unabhängig von der Identität des ersten Genomnukleotids und von Extensionen (+2 und +3 Nt) des 3'-Genomendes erfolgt (Deflubé et al., 2019). Dies konnte in der vorliegenden Arbeit ebenfalls bestätigt werden. Eine Deletion des ersten Genomnukleotids wirkte sich in den von uns verwendeten Replikations-kompetenten MG-Systemen nicht nachteilig im Vergleich zum Wildtyp-MG aus. Eine Extension des 3'-Endes um ein Nt (+G) steigerte die RNA-Synthese sämtlicher viraler RNA-Spezies sogar. Darüber hinaus zeigen die Daten, dass neben cRNA auch abortive *leader*RNAs spezifisch an C<sub>-2</sub> initiiert werden. Die Initiation der *leader*RNA-Synthese ist mechanistisch unabhängig von VP30. In Abwesenheit von VP30 sind die *leader*RNA-Spiegel erhöht. (NP) mRNA und *leader*RNA liegen 24 h nach EBOV-Infektion in einem Konzentrationsverhältnis von etwa 9:1 vor. Diese Aspekte sind mit der Hypothese konsistent, dass *leader*RNAs möglicherweise vorzeitig abgebrochene cRNAs und nicht zwangsläufig Vorlauf-Produkte der viralen Transkription darstellen. Die *leader*RNA-Synthese wird schwerpunktmäßig in einem Längenbereich von ca. 64 bis 79 Nt, d.h., im Bereich des NP-Hairpins, heterogen terminiert. Allerdings sind weder die Ausbildung der NP-Hairpin-Struktur noch alternative Terminationssignale oder die Präsenz von VP30 für diese Polymerisationsabbrüche verantwortlich.

Die Initiation der mRNA-Synthese an C<sub>-56</sub> kann prinzipiell VP30-unabhängig erfolgen. Eine effiziente Transkription findet jedoch nur unter Einbindung von VP30 statt, während jegliche Polymerisationsinitiation am terminalen 3'-Genomende (cRNA und *leader*RNA-Synthese) effizient ohne Beteiligung des Transkriptionsaktivators VP30 stattfinden kann. VP30 scheint folglich erst bei Initiation der NP mRNA-Transkription am TSS mit dem Polymerasekomplex zu assoziieren. Frühere Studien postulierten auf Basis relativ weniger experimenteller Daten die Relevanz der Sekundärstruktur im NP-TSS-Spacer-Bereich für die virale Transkriptionsregulation durch VP30 sowie für die Interaktion von NP und VP30 (Schlereth et al., 2016; Weik et al., 2002; Xu et al., 2017). Im Rahmen der vorliegenden Arbeit konnte erstmalig (indirekt) gezeigt werden, dass sich in der Tat RNA-Sekundärstrukturen während der Virus-Replikation und -Transkription ausbilden und die Feinregulation von Transkription und Replikation durch VP30 beeinflussen. In

monocistronischen Minigenomen erhöhten Stabilisierungen des TSS/Spacer-Hairpins die VP30-Abhängigkeit der Transkription, senkten jedoch – abhängig vom Ausmaß der Stabilisierung – die allgemeine virale RNA-Synthese z.T. dramatisch bis hin zu einer vollständigen Inaktivierung. Destabilisierungen des TSS/Spacer-Hairpins schwächten die strikte VP30-Abhängigkeit der viralen Transkription. Dies gilt allerdings für sämtliche Abweichungen von der authentischen Struktur des NP-Hairpins (Sequenz, Stabilität und Länge). Besonders Abweichungen in Sequenz und/oder Struktur des unteren Hairpin-Stammbereichs scheinen eine Relaxation der VP30-Abhängigkeit zu begünstigen. Möglicherweise spielt auch die Sequenz des hochkonservierten Genstart (GS)-Signals eine Rolle. Dieses weicht ausschließlich im Falle der NP- und L-Gene in der dritten Position von den GS-Sequenzen der anderen fünf EBOV-Gene ab (C statt A). Unabhängig davon lassen die Ergebnisse der vorliegenden Arbeit darauf schließen, dass die Ausbildung einer Sekundärstruktur am TSS keinesfalls obligat für die Virustranskription ist. Eine Deletion fast des gesamten Spacers (-12 Nt; ausgenommen G<sub>80</sub>), die simultan jegliche Sekundärstruktur unterbindet, reduzierte die Synthese sämtlicher viraler RNA-Spezies (vRNA, cRNA, *leader*RNA, mRNA) im Replikationskompetenten Minigenomsystem nur leicht bis moderat. Am stärksten beeinträchtigt wurde die mRNA-Synthese (ca. 55 % im Vergleich zum Wildtyp-3'-Leader). Trotz Aufhebung der NP-Hairpin-Struktur blieb die virale Transkription stark VP30-abhängig. Ein kontinuierliches UN<sub>5</sub>-Raster von U<sub>51</sub> - U<sub>128</sub> durch Einführung einer U<sub>75</sub>-Punktmutation steigerte die allgemeine virale RNA-Synthese weitgehend sekundärstrukturunabhängig, korrumpierte jedoch die Feinregulation von Transkription und Replikation durch VP30 (kaum oder keine Replikationsstimulation in Abwesenheit von VP30).

Darüber hinaus konnten wir zeigen, dass nicht nur cRNA- (Weik et al., 2005), sondern auch *leader*RNA- und mRNA-Synthese ausschließlich bei Einhaltung der Spacer-„Rule of 6“ (Hexamerphase) initiiert werden können. D.h. nur Spacer-Insertionen und -Deletionen eines Vielfachen der Zahl 6 wurden toleriert. Wie oben bereits erwähnt, hatte eine Deletion des Spacers um 12 Nt lediglich geringe Auswirkungen auf die allgemeine RNA-Synthese. Zudem zeigten Konstrukte mit Insertionen von bis zu 66 Nt im Spacer-Bereich immer noch eine ~ 25 %-ige Replikations-/Transkriptionsaktivität in Reporterassays. Diese Beobachtungen sind kaum mit der in der Literatur favorisierten Hypothese vereinbar, dass PE1 und PE2 während der RNA-Synthese auf derselben Seite der NP-RNA-Helix liegen müssen, um simultan vom selben Polymerasekomplex kontaktiert zu werden (Le Mercier and Kolakofsky, 2019; Vgl. Kapitel 1.8.1.3; Abb. 11). Wie im Falle der maximalen Spacer-Verkürzung (12 Nt-Deletion;  $\Delta 5'$  spacer-Konstrukt) beeinflussten Spacer-Verlängerungen die VP30-Abhängigkeit nur marginal.

Es ergibt sich somit folgendes Modell der Polymerisationsinitiation am 3'-Genomende: Jegliche Polymerisation wird offensichtlich vom zweiten 3'-terminalen Genomnukleotid aus initiiert und führt zur Synthese von *leaderRNAs*, die entweder zu Vollängen-Antigenomen elongiert oder im Bereich des ersten NP-Transkriptionsstarts abgebrochen werden. Grundvoraussetzung der Polymerisation dürfte dabei der Erhalt einer spezifischen Hexamerphase im Bereich des Spacers sein. Diese Hexamerperiodizität reguliert möglicherweise die NP-Enkapsidierung (6 Nt der RNA werden pro NP-Molekül gebunden; Sugita et al., 2018; Wan et al., 2017) und gewährleistet auf diese Weise, dass Replikationspromotorsignale und/oder *cis*-aktive Transkriptionssignale am 3'-Genomende vom Polymerasekomplex optimal erkannt werden können. Unklar ist, ob eine *leaderRNA*-Synthese Grundvoraussetzung für eine Polymerisations(re)initiation am TSS ist, oder ob lediglich ein initialer Kontakt des Polymerasekomplexes zum ersten Genomnukleotid und TSS stattfinden muss, der es dem RdRp-Komplex ermöglicht, ohne RNA-Synthese über den Leader zu scannen, bis das erste TSS erreicht wird. Am TSS bildet sich anschließend ein stabiler Transkriptionskomplex unter Einbindung von VP30 aus, der eine produktive Initiation und -Elongation der NP mRNA-Transkription ermöglicht. Möglicherweise wird die Einbindung von VP30 über die NP-Hairpin-Struktur reguliert. Letztere scheint hinsichtlich Struktur, Länge und Sequenz für ein ausgewogenes Verhältnis von allgemein effizienter RNA-Synthese und VP30-Abhängigkeit optimiert zu sein. Es ergibt sich somit ein Gesamtbild, nach dem das Virus über die VP30-Abhängigkeit sicherstellt, dass virale Transkription und Replikation, je nach Bedarf in den verschiedenen Infektionsphasen, dynamisch (über VP30-Phosphorylierung und -Dephosphorylierung) und sehr spezifisch reguliert werden.

## 6 AUSBLICK

Unter anderem deuten die Minigenom-Studien darauf hin, dass PE1 nicht nur den Replikationspromotor, sondern vermutlich auch den Transkriptionspromotor enthält. Es kann allerdings nach wie vor nicht ausgeschlossen werden, dass die PE1-Sequenz an sich gar nicht transkriptionsrelevant ist (im Sinne eines Transkriptionspromotors), sondern vielmehr (i) die Distanz des TSS-Signals zum 3'-Genomende (55 Nt) oder (ii) die exakte Position (= NP-Phase) des ersten TSS-Nukleotids C<sub>-56</sub> in der Bindetasche des entsprechenden NP-Moleküls relevant ist (Abb. 6). Um diese Fragen zu beantworten, könnte in Nachfolgestudien eine systematische Mutationsanalyse des 3'-Leader-Promotors im PE1-Bereich durchgeführt werden. Zunächst sollte geklärt werden, ob EBOV, ähnlich wie es für das Pneumovirus RSV gezeigt wurde, einen Replikations- und Transkriptionskernpromotor am 3'-Ende des Genoms besitzt (Noton and Fearn, 2015). In RSV, dessen Kernpromotor die 3'-terminalen 12 Nt umfasst, ähneln diese 12 Nt dem Startsignal des L-Gens. Sequenzanalysen der nah verwandten *Paramyxoviridae* (u.a. SeV) lassen darüber hinaus vermuten, dass auch hier eine Genstart-ähnliche Sequenz innerhalb der konservierten 3'-terminalen 12 Leader-Nukleotide vorliegt (Noton and Fearn, 2015). Folglich bieten sich für die Untersuchung des EBOV-3'-Leader-Promotors Deletionen von 6 und 12 Nukleotiden am 3'-Genomende an, die in Replikations-kompetenten/Transkriptions-defizienten Minigenomen hinsichtlich ihrer Replikations-relevanz überprüft werden können. Die parallele Verwendung Transkriptions-kompetenter/Replikations-defizienter Minigenome lässt Rückschlüsse zur Relevanz dieser Nukleotide für die virale Transkription zu. Hexamer-Deletionen von 6 oder 12 Nt stellen sicher, dass die NP-Phase [Punkt ii)] nicht verändert wird. Interessant wäre auch eine schrittweise Insertion von 6, 12 oder mehr Nukleotiden (ggf. bis hin zu 90, 180, 360, etc.) unmittelbar vor dem (3' vom) NP-TSS, ähnlich wie es für das Paramyxovirus SeV durchgeführt wurde (Vulliémoz et al., 2005). Auf diese Weise könnte überprüft werden, ob eine vergrößerte Distanz vom NP-TSS zum 3'-Genomende einen Einfluss auf die Transkriptionseffizienz am TSS hat. Für SeV wird angenommen, dass die konkurrierenden Transkriptions- und Replikationsfunktionen u.a. von der Replikationspromotorstärke und der Entfernung des ersten GS-Signals zum 3'-terminalen Genomende abhängig sind (Cordey and Roux, 2006; Vulliémoz et al., 2005). Im Fall von SeV scheint eine Positionierung des TSS weiter vom 3'-Ende entfernt die mRNA-Synthese proportional zur Distanz vom 3'-Ende zu verringern. Es wird postuliert, dass ein initialer Kontakt der Polymerase zum ersten Genomnukleotid besteht, welcher auch für die Distanzmessung zum ersten TSS relevant sein könnte (Cordey and Roux, 2006, 2007). Darüber hinaus gibt es Evidenzen, dass eine basale SeV-Transkription (~ 20 %) auch noch nach (fast) vollständiger Deletion von PE1 und/oder

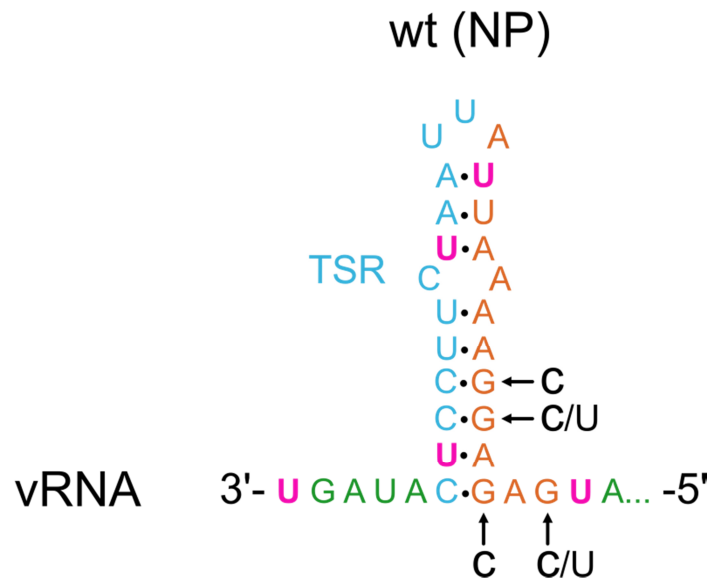


PE2 erfolgen kann. Einzige Voraussetzung ist, dass die PE1-Nukleotide unmittelbar vor dem TSS (Nt -49 bis -55) erhalten bleiben (Vulli  moz et al., 2005). Daraus ergibt sich ein transkriptionsrelevantes Sequenzelement zusammen mit dem TSS-Spacer-Bereich (Nt -56 bis -78) von in der Summe exakt 30 Nukleotiden. In diesem Sequenzelement liegt das erste TSS-Nukleotid in der Hexamerphasen-Position 2 (vom 3'-Ende aus gemessen), d.h., es liegt an der zweiten Position eines vom Nukleoprotein bedeckten RNA-Hexamers (Cordey and Roux, 2006). Es sei anzumerken, dass in diesen Studien Minigenome verwendet wurden, bei denen an das 3'-Ende des genomischen 3'-Leader-Promotors ein zus  tzlicher antigenomischer Promotor fusioniert wurde, der die Replikation dieser Minigenome steuert. Der genomische, replikationsinaktivierte Transkriptionspromotor befand sich dadurch 90 Nukleotide weiter intern. Die PE1/PE2-Deletionen und ihre Effekte auf die Transkription wurden an diesem internen Promotor studiert. Ein   hnlicher experimenteller Aufbau w  re auch f  r EBOV m  glich. Auf diese Weise k  nnte   berpr  ft werden, welche NP-Phase in diesem Fall f  r das erste TSS-Nukleotid relevant ist. Dar  ber hinaus kann in diesem Zusammenhang die Transkriptionsrelevanz des UN<sub>5</sub>-Hexamerrasters in der Region -51 bis -80 (in Summe 30 Nt) intensiver untersucht werden. Es bleibt zu kl  ren, ob diese Region f  r eine Replikationspromotor-unabh  ngige Transkription ausreichend ist. Man k  nnte zus  tzlich eine MG-Variante herstellen, in der das gesamte PE1 (inklusive Nt -51 bis -55) deletiert ist. So lie  e sich erkennen, ob das TSS allein f  r die Transkriptionsinitiation ausreichend ist.

In der vorliegenden Arbeit konnte zudem erstmalig – wenn auch indirekt – gezeigt werden, dass sich RNA-Sekund  rstrukturen tats  chlich w  hrend der Replikation und Transkription ausbilden und beide Prozesse beeinflussen. Dabei korrelierten die beobachteten Effekte durch RNA-Strukturstabilisierungen besser mit Ver  nderungen auf antigenomischer/mRNA-Ebene als auf genomischer vRNA-Ebene (siehe Manuskript 2; Kapitel 4.2). Die Regulation der viralen Polymerisation durch VP30 scheint, zumindest am ersten GS-Signal, bereits durch geringe Abweichungen von der NP-Hairpin-Sequenz, -Stabilit  t und -L  nge beeinflusst zu werden (Manuskript 4; Kapitel 4.4). Besonders Abweichungen in Sequenz und/oder Struktur des unteren Hairpin-Stammbereichs beg  nstigen eine Relaxierung der VP30-Abh  ngigkeit. Auff  llig ist, dass sowohl das MG-Konstrukt mit dem NP-Hairpin als auch MG-Konstrukte mit dem L-5'-UTR-Hairpin eine besonders strikte VP30-Abh  ngigkeit zeigen. Die Gemeinsamkeit beider Hairpinstrukturen ist, dass sie eine in der Sequenz leicht ver  nderte Genstart-Sequenz (C statt A an Position 3; siehe Tab. 1) im Vergleich zu den anderen EBOV-Genen aufweisen. M  glicherweise ist dieses C<sub>-58</sub> am TSS f  r die VP30-Abh  ngigkeit relevant. Dies k  nnte durch Substitution des C<sub>-58</sub>:G-Basenpaars mit einem A:U-

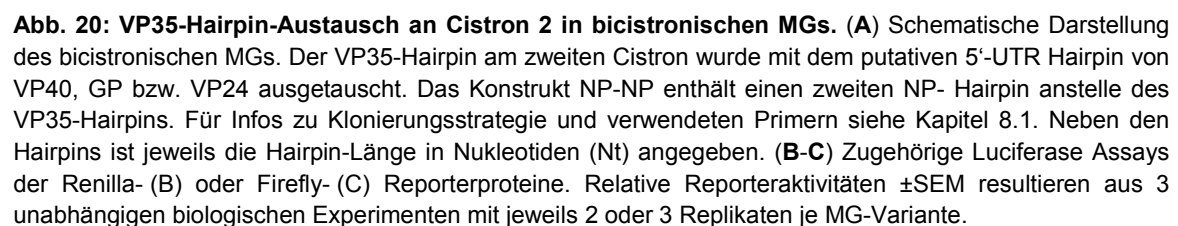
Basenpaar geklärt werden. Parallel ließe sich z.B. eine VP40-Variante vergleichend analysieren, in der umgekehrt das A:U-Basenpaar an dieser Position durch ein C:G-Basenpaar ausgetauscht wurde. Sollte die obige Hypothese zutreffen, so wäre eine relaxierte VP30-Abhängigkeit der eben beschriebenen NP-Mutante bzw. eine erhöhte VP30-Abhängigkeit des letztgenannten VP40-Konstrukts zu erwarten. Interessant ist in diesem Zusammenhang, dass ausgerechnet die NP- und L-Genstart-Sequenzen Erkennungssignale für die Interaktion mit zellulären Typ 1 Topoisomerasen enthalten. Die DNA-Topoisomerase I interagiert gemäß Takahashi *et al.* direkt mit EBOV L und steigert so die virale Transkription und Replikation (Takahashi et al., 2013). Sollte die dritte Position der Genstart-Sequenz tatsächlich VP30-relevant sein, böte sich auch die Untersuchung der Frage an, ob es möglicherweise zu einer RNA-vermittelten Interaktion zwischen VP30 und der DNA-Topoisomerase I kommt.

Die MG-Varianten der vorliegenden Arbeit ließen nur ungefähre Rückschlüsse zum Einfluss von destabilisierten TSS-Hairpinstrukturen auf die (VP30-abhängige) Transkription zu. In der NheI NP-Variante wurde die NP-Hairpin-Struktur destabilisiert, jedoch gleichzeitig auch die Sequenz im Spacer-Bereich stark verändert (6 Punktmutationen). In der NP U<sub>75</sub>-Variante wurde parallel zur Destabilisierung auch das UN<sub>5</sub>-Raster der TSS-Spacer-Region verändert. Darüber hinaus sind laut RNAfold-Vorhersagen sowohl beim NheI NP- als auch beim NP U<sub>75</sub>-Konstrukt Nukleotide der Promoterelemente PE1 und PE2 mit in die Bildung der Haarnadel-Helix auf vRNA-Ebene involviert. Auch dies könnte die VP30-Abhängigkeit beeinflusst haben. Es sei jedoch anzumerken, dass die experimentelle Ausrichtung der vorliegenden Arbeit nicht dahingehend konzipiert war, die Ausbildung von RNA-Sekundärstrukturen auf genomischer vRNA-Ebene *in vivo* zu verifizieren. Es bleibt also offen, ob Sekundärstrukturen nur auf mRNA- und ggf. antigenomischer cRNA-Ebene relevant sind (siehe oben). Um der Beantwortung dieser Frage näher zu kommen, könnten stabilisierte Hairpin-Varianten aus der vorliegenden Arbeit zusätzlich in Replikations-kompetenten/Transkriptions-defizienten Minigenomsystemen untersucht werden. In der Δ5' spacer-Mutante wurde zwar jegliche Sekundärstrukturausbildung unterbunden, jedoch simultan das Spacing zu PE2 verkürzt. Aus diesen Gründen ist es zielführend, eine Minigenomvariante ohne Struktur aber unter Erhalt der NP-Spacerlänge herzustellen. Es bieten sich die in Abb. 19 gezeigten Vierfach-Mutationen an, die gewährleisten, dass weder auf vRNA- noch auf mRNA-Ebene Sekundärstrukturen ausgebildet werden können (RNAfold-Strukturvorhersage).

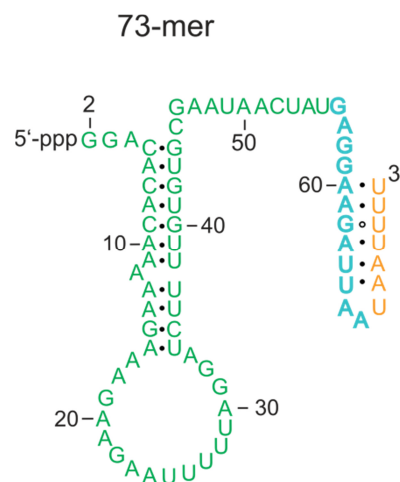


**Abb. 19: NP-Spacer-Variante mit aufgelöster Sekundärstruktur auf vRNA- und mRNA-Ebene aber ohne Veränderung der Spacer-Länge.** Die Sekundärstrukturvorhersage mittels RNAfold erfolgte unter Verwendung von Standard-Parametern (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>).

Während die Hexamerphase im TSS-Bereich jegliche Polymerisation am 3'-Genomende steuert (cRNA-, leaderRNA-, und mRNA-Synthese), gibt es erste Hinweise, dass sie intern keine Rolle spielt (Abb. 20). Bicistronische Minigenome, in denen der putative VP35-5'-UTR-Hairpin am zweiten Cistron durch den NP-Hairpin oder den putativen VP40-, GP- oder VP24-5'-UTR-Hairpins ausgetauscht wurde, zeigten keine nennenswerten Transkriptionseffizienzunterschiede zum Wildtyp (NP-VP35)-MG (Abb. 20). Dies erscheint logisch, da intern keine Replikationspromotoren zu finden sind, die mit Transkriptions-relevanten Sequenzen überlappen. Die Daten deuten zudem darauf hin, dass Hairpinstruktur- und Längenvariationen an internen Genstarts besser toleriert werden. Sicherlich sollten diesbezüglich weitere Mutationsanalysen in Nachfolgestudien durchgeführt werden. Es bieten sich bicistronische MG-Varianten mit dem VP40+30 nt-Hairpin (insg. 66 Nt länger als der NP-Hairpin) sowie mit den stabilisierten NP- und VP40-Varianten (siehe Manuskript 2; Kapitel 4.2) an. Ebenso könnte die in Abb. 19 aufgeführte NP-Spacer-Mutante mit nativer Spacer-Länge (13 Nt) aber ohne 2D-Struktur am internen Genstart analysiert werden. Ein weiterer interessanter Punkt ist, dass alle in Abb. 20 aufgeführten bicistronischen MG-Varianten sowohl am ersten als auch intern am zweiten Cistron eine strikt VP30-abhängige Transkription zeigen. Eine Studie aus dem Jahr 2002 stellte die Hypothese auf, dass der NP-Hairpin am ersten TSS die VP30-Abhängigkeit der Transkription aller internen mRNAs vorgibt. Die Autoren konnten die Transkription am zweiten Cistron eines bicistronischen MGs (partiell) VP30-unabhängig machen, indem sie den NP-Hairpin am ersten Cistron durch den (destabilisierten) NheI NP-Hairpin ersetzten (Weik et al., 2002). In zukünftigen Studien könnte die VP30-



Schließlich könnte die mögliche Funktion der Leader-Transkripte im Kontext der viralen Infektion weiter untersucht werden. So konnte im Rahmen dieser Arbeit nicht abschließend geklärt werden, ob die *leader*RNAs eine Rolle in der angeborenen Immunantwort spielen. Bislang wurde für T7-transkribierte EBOV-5'-PPP-*leader*RNAs (Nt 1 – 55; (Louber et al., 2014)) und synthetische 5'-PPP-*leader*RNAs (Nt 1 – 65; vorläufige Daten, die nicht in diese Arbeit aufgenommen wurden) eine nur geringe bis keine Induktion der RIG-I-vermittelten Immunantwort festgestellt. Allerdings enthielten die verwendeten *leader*RNAs in beiden Fällen das erste Genomnukleotid C<sub>1</sub>. Es ist daher sinnvoll, die Experimente mit einem authentischen Leader-Transkript (z.B. Nt G<sub>2</sub> – U<sub>73</sub>; Abb. 21) zu wiederholen. Dies ist insbesondere deshalb naheliegend, weil frühere Studien zur RIG-I-vermittelten Immunantwort Hinweise lieferten, dass ein 5'-PPP-C-Nukleotid möglicherweise ungünstig für eine RIG-I-Erkennung ist, während 5'-PPP-Purin-dsRNAs bevorzugt erkannt werden (Schlee et al., 2009; Schmidt et al., 2009). Allerdings lassen RNA-Sekundärstrukturvorhersagen und Struktur-Probing der *leader*RNA (Supplementary Information aus Schlereth et al., 2016) vermuten, dass an G<sub>2</sub> gestartete Leader-Transkripte aller Voraussicht nach 5'-Überhänge von drei Nukleotiden aufweisen (siehe Abb. 21), die auf Grundlage der o.g. Literatur eher ungünstig für die RIG-I-Erkennung sein sollten.



**Abb. 21: Putative Sekundärstruktur eines Leader-Transkripts (leaderRNA; 73 Nt);** RNAfold-Struktur-vorhersage, verifiziert durch Struktur-Probing (Supplementary Information, Schlereth et al., 2016).

Auf Basis der experimentellen Daten dieser Arbeit könnten die *leader*RNA-Mengen sowohl nach EBOV-Infektion als auch nach MG-Transfektion wesentlich geringer sein als die der ersten mRNA (NP mRNA bzw. Renilla mRNA). Eine mögliche Erklärung ist, dass *leader*RNAs abortive cRNAs und keine obligaten Vorlauf-Produkte der mRNA-Synthese sind. Andererseits ließ sich im Rahmen der experimentellen Möglichkeiten dieser Arbeit nicht klären, ob die *leader*RNAs in den Zellen möglicherweise schneller als

die mRNAs abgebaut wurden, da sie sehr wahrscheinlich weder 5'-Kappe und Poly(A)-Schwanz tragen noch durch NP enkapsidiert werden. Immerhin gibt es einige Literaturbeispiele, in denen im Kontext anderer NNS-Virus-Leader-Transkripte von Zellsystem-abhängigen Unterschieden in ihren Abbauraten berichtet wurde (Horikami and Moyer, 1991; Vidal and Kolakofsky, 1989). Die Haupt-5'-Exonukleasen in eukaryotischen Zellen sind Xrn1 (hauptsächlich zytoplasmatisch) und Xrn2 (hauptsächlich nukleär). Die Aktivitäten beider Enzyme werden zunächst durch die Anwesenheit von 5'-Triphosphatgruppen (5'-PPP) gehemmt. 5'-PPP-Enden werden üblicherweise vom nukleären Exosom abgebaut, einem 3' – 5'-Exonukleasekomplex mit zwei katalytischen Untereinheiten (Rrp6 und Dis3). Dis3 besitzt zusätzlich eine endonukleolytische Aktivität. Die durch den Dis3-/Rrp6-Abbau neu generierten 5'-Monophosphat-Enden können anschließend von Xrn1 und Xrn2 weiter abgebaut werden. Das Exosom ist am Abbau einer Vielzahl zellulärer Transkripte beteiligt; rRNA, kleine nukleäre und nukleoläre RNAs, fehlerhafte Prä-rRNAs, Prä-tRNAs, mRNAs oder Produkte pervasiver Transkription (nicht-kodierende RNAs) [Zusammengefasst in (Bresson and Tollervey, 2018)]. Da die EBOV-Replikation und -Transkription in Einschlusskörperchen im Zytoplasma stattfinden, bleibt unklar, inwieweit nukleäre Exo- und Endonukleasen am Abbau der viralen *leader* RNAs beteiligt sind. Auch dürften die EBOV-*leader* RNAs zunächst vor Abbau durch die zytoplasmatische 5'-Exonuklease Xrn1 geschützt sein, wenn sie wie angenommen ein 5'-PPP-Ende besitzen. Den *leader* RNA-Turnover könnte man in Nachfolgestudien zunächst mit einem Xrn1-Einzel-Knockout analysieren. Je nach Ergebnis könnten dann ergänzend Knockouts von Dis3 und ggf. Xrn2 folgen. Bevor jedoch aufwendige Knockout-Experimente durchgeführt werden, sollte zunächst geklärt werden, ob die Art der RNA-Extraktion oder -Aufreinigung nach DNase-Verdau eine Unterrepräsentation der *leader* RNAs verursacht haben könnte. Für Gesamt-RNA-Extraktionen aus EBOV-infizierten HuH7-Zellen wurde das *mirVana*<sup>™</sup> miRNA Isolation Kit (+ Phenol; Thermo Fisher Scientific) verwendet, welches gemäß Herstellerangaben Gesamt-RNA-Extraktionen einschließlich kleinerer RNAs < 200 Nt ermöglicht. Für RNA-Extraktionen aus MG-transfizierten HEK293-Zellen wurde das RNeasy Mini Kit (QIAGEN) verwendet, inklusive eines zweifachen DNase-Verdaus (1x Säulen-basiert und 1x in Lösung mit anschließender Phenol-Chloroform-Isoamylalkohol-Extraktion). Das RNeasy Mini Kit eignet sich laut Hersteller bevorzugt für Extraktionen von RNAs > 200 Nt, eine Unterrepräsentation der *leader* RNAs in der Gesamt-RNA liegt hier folglich im Bereich des Möglichen. Dies könnte auch der Grund für die abweichenden Mengen an Leader-Transkripten im Vergleich von EBOV-infizierten Zellen (9-fach geringer als NP mRNA) zu MG-transfizierten Zellen (68-fach geringer als Renilla mRNA) sein. Um Unterschiede in der Repräsentierung von RNAs

verschiedener Längenbereiche für die jeweilige Extraktionsmethode zu identifizieren, müsste das Zelllysate mit (synthetischen oder T7-transkribierten) Kontroll-RNAs in der Länge von *leader*RNA und mRNA versetzt werden und anschließend ihre Wiedergewinnung (*Recovery*) nach Aufreinigung über die Säulenmatrix überprüft werden (mittels qRT-PCR). Hier könnte die bereits verwendete synthetische *leader*RNA aus dieser Arbeit verwendet werden. Eine Referenz-mRNA müsste allerdings relativ aufwendig mittels T7-Transkription und anschließender (z.B. säulenchromatografischer) Aufreinigung hergestellt werden.

Darüber hinaus könnte mit Hilfe von iVLP-Systemen (Kapitel 1.9; Abb. 13) untersucht werden, ob Leader-Transkripte unter Primärtranskriptionsbedingungen gebildet werden. Sollten *leader*RNAs tatsächlich abgebrochene cRNAs sein, so sind in diesem Fall keine Leader-Transkripte zu erwarten. In iVLP-Assays kann über die Infektion von naiven Zellen (= Indikatorzellen) mit infektiösen, virusartigen Partikeln (iVLPs) eine EBOV-Infektion simuliert werden. Da jedoch ebenfalls ein Minigenom anstelle eines authentischen EBOV-Genoms verwendet wird, in dem die EBOV-Gene durch ein Reporterprotein ersetzt sind, ist der virale Lebenszyklus in diesen naiven Indikatorzellen auf die Primärtranskription beschränkt. Im Anschluss an das iVLP-Assay können die Leader-Transkriptmengen in der Gesamt-RNA der Indikatorzellen mittels qRT-PCR evaluiert werden. Erste entsprechende Experimente wurden bereits im Rahmen dieser Arbeit durchgeführt (Daten nicht gezeigt), allerdings offenbarten diese, dass die Verwendung eines klassischen monocistronischen Minigenoms keine ausreichende Infektiosität der iVLPs und damit keine ausreichenden Gesamt-RNA-Mengen liefert. Entsprechend konnten mittels qRT-PCR der Indikatorzell-Gesamt-RNA keine validen Quantifizierungen vorgenommen werden. Zukünftig sollten daher tetracistronische MG-Systeme eingesetzt werden. Tetracistronische Minigenome kodieren nicht nur das Reporterprotein sondern zusätzlich die EBOV-Gene VP40, GP<sub>1,2</sub> und VP24. Genau wie beim klassischen VLP-Assay liefert das tetracistronische VLP-System infektiöse VLPs, die genutzt werden können, um naive Zellen zu infizieren und die Primärtranskription zu analysieren. Der entscheidende Vorteil dieses Systems ist, dass die Gene von VP40, GP<sub>1,2</sub> und VP24 während der viralen Transkription in den sogenannten *iVLP-Producer-Zellen* "normal" exprimiert und nicht über Plasmide überexprimiert werden. In der Konsequenz führt dies zu Kinetiken und Expressions-Spiegeln dieser Proteine, die dem viralen Lebenszyklus wesentlich mehr ähneln und daher auch zu ca. 500-fach höheren Mengen an infektiösen VLPs führen als bei klassischen iVLP-Systemen, die auf monocistronischen Minigenomen basieren (Hoenen et al., 2014; Watt et al., 2014).

## 7 LITERATURVERZEICHNIS

- Albariño, C. G., Guerrero, L. W., Chakrabarti, A. K., and Nichol, S. T. (2018). Transcriptional analysis of viral mRNAs reveals common transcription patterns in cells infected by five different filoviruses. *PLoS One* 13, 1–13. doi:10.1371/journal.pone.0201827.
- Albertini, A. A. V. (2006). Crystal Structure of the Rabies Virus Nucleoprotein-RNA Complex. *Science* (80-. ). 313, 360–363. doi:10.1126/science.1125280.
- Amarasinghe, G. K., Ayllón, M. A., Bào, Y., Basler, C. F., Bavari, S., Blasdel, K. R., et al. (2019). Taxonomy of the order Mononegavirales: update 2019. *Arch. Virol.* 164, 1967–1980. doi:10.1007/s00705-019-04247-4.
- Banadyga, L., Hoenen, T., Ambroggio, X., Dunham, E., Groseth, A., and Ebihara, H. (2017). Ebola virus VP24 interacts with NP to facilitate nucleocapsid assembly and genome packaging. *Sci. Rep.* 7, 7698. doi:10.1038/s41598-017-08167-8.
- Bào, Y., Amarasinghe, G. K., Basler, C. F., Bavari, S., Bukreyev, A., Chandran, K., et al. (2017). Implementation of Objective PASC-Derived Taxon Demarcation Criteria for Official Classification of Filoviruses. *Viruses* 9. doi:10.3390/v9050106.
- Barr, J. N., and Wertz, G. W. (2004). Bunyamwera Bunyavirus RNA Synthesis Requires Cooperation of 3'- and 5'-Terminal Sequences. *J. Virol.* 78, 1129–1138. doi:10.1128/jvi.78.3.1129-1138.2004.
- Basler, C. F. (2017). Molecular pathogenesis of viral hemorrhagic fever. *Semin. Immunopathol.* 39, 551–561. doi:10.1007/s00281-017-0637-x.
- Biedenkopf, N., Hartlieb, B., Hoenen, T., and Becker, S. (2013). Phosphorylation of Ebola Virus VP30 Influences the Composition of the Viral Nucleocapsid Complex. *J. Biol. Chem.* 288, 11165–11174. doi:10.1074/jbc.m113.461285.
- Biedenkopf, N., and Hoenen, T. (2017). “Modeling the Ebolavirus Life Cycle with Transcription and Replication-Competent Viruslike Particle Assays,” in *Methods in molecular biology* (Clifton, N.J.), 119–131. doi:10.1007/978-1-4939-7116-9\_9.
- Biedenkopf, N., Lier, C., and Becker, S. (2016a). Dynamic Phosphorylation of VP30 Is Essential for Ebola Virus Life Cycle. *J. Virol.* 90, 4914–4925. doi:10.1128/jvi.03257-15.
- Biedenkopf, N., Schlereth, J., Grünweller, A., Becker, S., and Hartmann, R. K. (2016b). RNA Binding of Ebola Virus VP30 Is Essential for Activating Viral Transcription. *J. Virol.* 90, 7481–7496. doi:10.1128/jvi.00271-16.
- Bitko, V., Musiyenko, A., Bayfield, M. A., Maraia, R. J., and Barik, S. (2008). Cellular La protein shields nonsegmented negative-strand RNA viral leader RNA from RIG-I and enhances virus growth by diverse mechanisms. *J. Virol.* 82, 7977–87. doi:10.1128/JVI.02762-07.
- Blakqori, G., Kochs, G., Haller, O., and Weber, F. (2003). Functional L polymerase of La Crosse virus allows in vivo reconstitution of recombinant nucleocapsids. *J. Gen. Virol.* 84, 1207–1214. doi:10.1099/vir.0.18876-0.
- Blumberg, B. M., Leppert, M., and Kolakofsky, D. (1981). Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication. *Cell* 23, 837–45. doi:10.1016/0092-8674(81)90448-7.
- Bosio, C. M., Aman, M. J., Grogan, C., Hogan, R., Ruthel, G., Negley, D., et al. (2003). Ebola and Marburg Viruses Replicate in Monocyte-Derived Dendritic Cells without Inducing the Production of Cytokines and Full Maturation. *J. Infect. Dis.* 188, 1630–1638. doi:10.1086/379199.



- Brauburger, K., Boehmann, Y., Kräling, V., and Mühlberger, E. (2016). Structure and Regulatory Elements on Gene Expression and Polymerase Scanning Behavior. *J. Virol.* 90, 1898–1909. doi:10.1128/JVI.02341-15.Editor.
- Brauburger, K., Boehmann, Y., Tsuda, Y., Hoenen, T., Olejnik, J., and Schümann, M. (2014a). Analysis of the Highly Diverse Gene Borders in Ebola Virus Reveals a Distinct Mechanism of Transcriptional Regulation. 88, 12558–12571. doi:10.1128/JVI.01863-14.
- Brauburger, K., Deflubé, L. R., and Mühlberger, E. (2014b). *Filovirus Transcription and Replication*. doi:10.1142/9789814635349\_0020.
- Brès, P. (1978). [The epidemic of Ebola haemorrhagic fever in Sudan and Zaire, 1976: introductory note]. *Bull. World Health Organ.* 56, 245. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/307454> [Accessed October 11, 2019].
- Bresson, S., and Tollervey, D. (2018). Surveillance-ready transcription: Nuclear RNA decay as a default fate. *Open Biol.* 8. doi:10.1098/rsob.170270.
- Briand, S., Bertherat, E., Cox, P., Formenty, P., Kieny, M.-P., Myhre, J. K., et al. (2014). The international Ebola emergency. *N. Engl. J. Med.* 371, 1180–3. doi:10.1056/NEJMp1409858.
- Bukreyev, A. A., Chandran, K., Dolnik, O., Dye, J. M., Ebihara, H., Leroy, E. M., et al. (2014). Discussions and decisions of the 2012–2014 International Committee on Taxonomy of Viruses (ICTV) Filoviridae Study Group, January 2012–June 2013. *Arch. Virol.* 159, 821–30. doi:10.1007/s00705-013-1846-9.
- Burk, R., Bollinger, L., Johnson, J. C., Wada, J., Radoshitzky, S. R., Palacios, G., et al. (2016). Neglected filoviruses. *FEMS Microbiol. Rev.* 40, 494–519. doi:10.1093/femsre/fuw010.
- Calain, P., Monroe, M. C., and Nichol, S. T. (1999). Ebola Virus Defective Interfering Particles and Persistent Infection. *Virology* 262, 114–128. doi:10.1006/VIRO.1999.9915.
- Calain, P., and Roux, L. (1993). The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *J. Virol.* 67, 4822–30. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8392616> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC237869>.
- Calain, P., and Roux, L. (1995). Functional characterisation of the genomic and antigenomic promoters of Sendai virus. *Virology*. doi:10.1006/viro.1995.1464.
- Chuang, J. L., and Perrault, J. (1997). Initiation of vesicular stomatitis virus mutant polR1 transcription internally at the N gene in vitro. *J. Virol.* 71, 1466–75. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8995672> [Accessed September 7, 2019].
- Commission, R. of an I. (1978). Ebola haemorrhagic fever in Zaire, 1976. *Bull. World Health Organ.* 56, 271–93. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/307456> [Accessed October 11, 2019].
- Conzelmann, K. K. (2004). “Reverse Genetics of Mononegavirales,” in (Springer, Berlin, Heidelberg), 1–41. doi:10.1007/978-3-662-06099-5\_1.
- Cordey, S., and Roux, L. (2006). Transcribing paramyxovirus RNA polymerase engages the template at its 3' extremity. *J. Gen. Virol.* 87, 665–672. doi:10.1099/vir.0.81353-0.
- Cordey, S., and Roux, L. (2007). Further characterization of a paramyxovirus transcription initiation signal: Search for required nucleotides upstream and importance of the N phase context. *J. Gen. Virol.* 88, 1555–1564. doi:10.1099/vir.0.82701-0.

- Cowton, V. M., and Fearn, R. (2005). Evidence that the respiratory syncytial virus polymerase is recruited to nucleotides 1 to 11 at the 3' end of the nucleocapsid and can scan to access internal signals. *J. Virol.* 79, 11311–22. doi:10.1128/JVI.79.17.11311-11322.2005.
- Cowton, V. M., McGivern, D. R., and Fearn, R. (2006). Unravelling the complexities of respiratory syncytial virus RNA synthesis. *J. Gen. Virol.* 87, 1805–1821. doi:10.1099/vir.0.81786-0.
- Cox, R., Pickar, A., Qiu, S., Tsao, J., Rodenburg, C., Dokland, T., et al. (2014). Structural studies on the authentic mumps virus nucleocapsid showing uncoiling by the phosphoprotein. *Proc. Natl. Acad. Sci.* 111, 15208–15213. doi:10.1073/pnas.1413268111.
- Crary, S. M., Towner, J. S., Honig, J. E., Shoemaker, T. R., and Nichol, S. T. (2003). Analysis of the role of predicted RNA secondary structures in Ebola virus replication. *Virology* 306, 210–218. doi:10.1016/S0042-6822(02)00014-4.
- Davey, R. A., Shtanko, O., Anantpadma, M., Sakurai, Y., Chandran, K., and Maury, W. (2017). “Mechanisms of filovirus entry,” in *Current Topics in Microbiology and Immunology* (Springer, Cham), 323–352. doi:10.1007/82\_2017\_14.
- Davey, R. T., Dodd, L., Proschan, M. A., Neaton, J., Nordwall, J. N., Koopmeiners, J. S., et al. (2016). A randomized, controlled trial of ZMapp for ebola virus infection. *N. Engl. J. Med.* 375, 1448–1456. doi:10.1056/NEJMoa1604330.
- Deen, G. F., Broutet, N., Xu, W., Knust, B., Sesay, F. R., McDonald, S. L. R., et al. (2017). Ebola RNA Persistence in Semen of Ebola Virus Disease Survivors — Final Report. *N. Engl. J. Med.* 377, 1428–1437. doi:10.1056/NEJMoa1511410.
- Deflubé, L. R., Cressey, T. N., Hume, A. J., Olejnik, J., Haddock, E., and Feldmann, F. (2019). Ebolavirus polymerase uses an unconventional genome replication mechanism. 116, 8535–8543. doi:10.1073/pnas.1815745116.
- Deng, I. M., Duku, O., and Gillo, A. L. (1978). Ebola haemorrhagic fever in Sudan, 1976. Report of a WHO/International Study Team. *Bull. World Health Organ.* 56, 247–270. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/307455> [Accessed October 11, 2019].
- Diallo, B., Sissoko, D., Loman, N. J., Bah, H. A., Bah, H., Worrell, M. C., et al. (2016). Resurgence of Ebola Virus Disease in Guinea Linked to a Survivor With Virus Persistence in Seminal Fluid for More Than 500 Days. *Clin. Infect. Dis.* 63, 1353–1356. doi:10.1093/cid/ciw601.
- Dolnik, O., Stevermann, L., Kolesnikova, L., and Becker, S. (2015). Marburg virus inclusions: A virus-induced microcompartment and interface to multivesicular bodies and the late endosomal compartment. *Eur. J. Cell Biol.* 94, 323–331. doi:10.1016/j.ejcb.2015.05.006.
- Egelman, E. H., Wu, S. S., Amrein, M., Portner, A., and Murti, G. (1989). The Sendai virus nucleocapsid exists in at least four different helical states. *J. Virol.* 63, 2233–43. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC250641/> [Accessed August 30, 2019].
- Enterlein, S., Schmidt, K. M., Schumann, M., Conrad, D., Krahling, V., Olejnik, J., et al. (2009). The Marburg Virus 3' Noncoding Region Structurally and Functionally Differs from That of Ebola Virus. *J. Virol.* 83, 4508–4519. doi:10.1128/jvi.02429-08.
- Fearn, R., Peeples, M. E., and Collins, P. L. (2002). Mapping the Transcription and Replication Promoters of Respiratory Syncytial Virus. *J. Virol.* 76, 1663–1672. doi:10.1128/jvi.76.4.1663-1672.2002.
- Fearn, R., and Plemper, R. K. (2017). Polymerases of paramyxoviruses and

- pneumoviruses. *Virus Res.* 234, 87–102. doi:10.1016/j.virusres.2017.01.008.
- Feldmann, H., Mühlberger, E., Randolph, A., Will, C., Kiley, M. P., Sanchez, A., et al. (1992). Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus Res.* 24, 1–19. doi:10.1016/0168-1702(92)90027-7.
- Flick, R., and Hobom, G. (1999). Interaction of influenza virus polymerase with viral RNA in the “corkscrew” conformation. *J. Gen. Virol.* doi:10.1099/0022-1317-80-10-2565.
- Fodor, E., Pritlove, D. C., and Brownlee, G. G. (1994). The influenza virus panhandle is involved in the initiation of transcription. *J. Virol.* 68, 4092–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8189550>.
- Funke, C., Becker, S., Dartsch, H., Klenk, H.-D., and Mühlberger, E. (1995). Acylation of the Marburg Virus Glycoprotein. *Virology* 208, 289–297. doi:10.1006/viro.1995.1151.
- Garcin, D., and Kolakofsky, D. (1992). Tacaribe arenavirus RNA synthesis in vitro is primer dependent and suggests an unusual model for the initiation of genome replication. *J. Virol.* 66, 1370–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1738196> [Accessed September 7, 2019].
- Garske, T., Cori, A., Ariyaratnam, A., Blake, I. M., Dorigatti, I., Eckmanns, T., et al. (2017). Heterogeneities in the case fatality ratio in the west African Ebola outbreak 2013–2016. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160308. doi:10.1098/rstb.2016.0308.
- Gaudinski, M. R., Coates, E. E., Novik, L., Widge, A., Houser, K. V., Burch, E., et al. (2019). Safety, tolerability, pharmacokinetics, and immunogenicity of the therapeutic monoclonal antibody mAb114 targeting Ebola virus glycoprotein (VRC 608): an open-label phase 1 study. *Lancet* 393, 889–898. doi:10.1016/S0140-6736(19)30036-4.
- Geisbert, T. W., and Jahrling, P. B. (1995). Differentiation of filoviruses by electron microscopy. *Virus Res.* 39, 129–150. doi:10.1016/0168-1702(95)00080-1.
- Ghosh, A., Nayak, R., and Shaila, M. S. (1996). Synthesis of leader RNA and editing of P mRNA during transcription by rinderpest virus. *Virus Res.* 41, 69–76. doi:10.1016/0168-1702(95)01276-1.
- Glynn, J. R., Bower, H., Johnson, S., Houlihan, C. F., Montesano, C., Scott, J. T., et al. (2017). Asymptomatic infection and unrecognised Ebola virus disease in Ebola-affected households in Sierra Leone: a cross-sectional study using a new non-invasive assay for antibodies to Ebola virus. *Lancet. Infect. Dis.* 17, 645–653. doi:10.1016/S1473-3099(17)30111-1.
- Goldstein, T., Anthony, S. J., Gbakima, A., Bird, B. H., Bangura, J., Tremeau-Bravard, A., et al. (2018). The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses. *Nat. Microbiol.* 3, 1084–1089. doi:10.1038/s41564-018-0227-2.
- Gordon, T. B., Hayward, J. A., Marsh, G. A., Baker, M. L., and Tachedjian, G. (2019). Host and viral proteins modulating ebola and marburg virus egress. *Viruses* 11. doi:10.3390/v11010025.
- Green, T. J., Zhang, X., Wartz, G. W., and Luo, M. (2006). Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* (80- ). 313, 357–360. doi:10.1126/science.1126953.
- Grünweller, A., and Hartmann, R. K. (2007). Locked Nucleic Acid Oligonucleotides. *BioDrugs* 21, 235–243. doi:10.2165/00063030-200721040-00004.
- Hanley, L. L., McGivern, D. R., Teng, M. N., Djang, R., Collins, P. L., and Fearn, R.

- (2010). Roles of the respiratory syncytial virus trailer region: effects of mutations on genome production and stress granule formation. *Virology* 406, 241–52. doi:10.1016/j.virol.2010.07.006.
- Hinzman, E. E., Barr, J. N., and Wertz, G. W. (2002). Identification of an upstream sequence element required for vesicular stomatitis virus mRNA transcription. *J. Virol.* 76, 7632–41. doi:10.1128/jvi.76.15.7632-7641.2002.
- Hoch, P. G., Schlereth, J., Lechner, M., and Hartmann, R. K. (2016). *Bacillus subtilis* 6S-2 RNA serves as a template for short transcripts in vivo. *RNA* 22, 614–622. doi:10.1261/rna.055616.115.
- Hoenen, T., Groseth, A., de Kok-Mercado, F., Kuhn, J. H., and Wahl-Jensen, V. (2011). Minigenomes, transcription and replication competent virus-like particles and beyond: Reverse genetics systems for filoviruses and other negative stranded hemorrhagic fever viruses. *Antiviral Res.* 91, 195–208. doi:10.1016/j.antiviral.2011.06.003.
- Hoenen, T., Groseth, A., and Feldmann, H. (2019). Therapeutic strategies to target the Ebola virus life cycle. *Nat. Rev. Microbiol.* 17, 593–606. doi:10.1038/s41579-019-0233-2.
- Hoenen, T., Groseth, A., Kolesnikova, L., Theriault, S., Ebihara, H., Hartlieb, B., et al. (2006). Infection of Naïve Target Cells with Virus-Like Particles: Implications for the Function of Ebola Virus VP24. *J. Virol.* 80, 7260–7264. doi:10.1128/JVI.00051-06.
- Hoenen, T., Jung, S., Herwig, A., Groseth, A., and Becker, S. (2010). Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription. *Virology* 403, 56–66. doi:10.1016/j.virol.2010.04.002.
- Hoenen, T., Shabman, R. S., Groseth, A., Herwig, A., Weber, M., Schudt, G., et al. (2012). Inclusion Bodies Are a Site of Ebolavirus Replication. *J. Virol.* 86, 11779–11788. doi:10.1128/jvi.01525-12.
- Hoenen, T., Watt, A., Mora, A., and Feldmann, H. (2014). Modeling the lifecycle of ebola virus under biosafety level 2 conditions with virus-like particles containing tetracistronic minigenomes. *J. Vis. Exp.*, 1–10. doi:10.3791/52381.
- Hoffman, M. A., and Banerjee, A. K. (2000a). Analysis of RNA secondary structure in replication of human parainfluenza virus type 3. *Virology*. doi:10.1006/viro.2000.0369.
- Hoffman, M. A., and Banerjee, A. K. (2000b). Precise Mapping of the Replication and Transcription Promoters of Human Parainfluenza Virus Type 3. *Virology* 269, 201–211. doi:10.1006/VIRO.2000.0223.
- Hoffmann, S., Otto, C., Kurtz, S., Sharma, C. M., Khaitovich, P., Vogel, J., et al. (2009). Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput. Biol.* 5, e1000502. doi:10.1371/journal.pcbi.1000502.
- Honda, T., Sofuku, K., Kojima, S., Yamamoto, Y., Ohtaki, N., and Tomonaga, K. (2017). Linkage between the leader sequence and leader RNA production in Borna disease virus-infected cells. *Virology*. doi:10.1016/j.virol.2017.07.011.
- Horikami, S. M., and Moyer, S. A. (1991). Synthesis of leader RNA and editing of the P mRNA during transcription by purified measles virus. *J. Virol.* 65, 5342–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1895388> [Accessed September 4, 2019].
- Hume, A. J., and Mühlberger, E. (2019). Distinct Genome Replication and Transcription Strategies within the Growing Filovirus Family. *J. Mol. Biol.* doi:10.1016/j.jmb.2019.06.029.
- Ilinykh, P. A., Tigabu, B., Ivanov, A., Ammosova, T., Obukhov, Y., Garron, T., et al.

- (2014). Role of protein phosphatase 1 in dephosphorylation of Ebola virus VP30 protein and its targeting for the inhibition of viral transcription. *J. Biol. Chem.* 289, 22723–38. doi:10.1074/jbc.M114.575050.
- Iseni, F., Garcin, D., Nishio, M., Kedersha, N., Anderson, P., and Kolakofsky, D. (2002). Sendai virus trailer RNA binds TIAR, a cellular protein involved in virus-induced apoptosis. *EMBO J.* 21, 5141. doi:10.1093/EMBOJ/CDF513.
- Ivanov, A., Ramanathan, P., Parry, C., Ilinykh, P. A., Lin, X., Petukhov, M., et al. (2019). Global phosphoproteomic analysis of Ebola virions reveals a novel role for VP35 phosphorylation-dependent regulation of genome transcription. *Cell. Mol. Life Sci.*, 1–25. doi:10.1007/s00018-019-03303-1.
- Johnson, K. M., Lange, J. V., Webb, P. A., and Murphy, F. A. (1977). ISOLATION AND PARTIAL CHARACTERISATION OF A NEW VIRUS CAUSING ACUTE HAEMORRHAGIC FEVER IN ZAIRE. *Lancet* 309, 569–571. doi:10.1016/S0140-6736(77)92000-1.
- Kemenesi, G., Kurucz, K., Dallos, B., Zana, B., Földes, F., Boldogh, S., et al. (2018). Re-emergence of Lloviu virus in *Miniopterus schreibersii* bats, Hungary, 2016. *Emerg. Microbes Infect.* 7, 1–4. doi:10.1038/s41426-018-0067-4.
- Kirchdoerfer, R. N., Abelson, D. M., Li, S., Wood, M. R., and Saphire, E. O. (2015). Assembly of the Ebola Virus Nucleoprotein from a Chaperoned VP35 Complex. *Cell Rep.* 12, 140–149. doi:10.1016/j.celrep.2015.06.003.
- Kirchdoerfer, R. N., Moyer, C. L., Abelson, D. M., and Saphire, O. (2016). The Ebola Virus VP30-NP Interaction Is a Regulator of Viral RNA Synthesis. 1–22. doi:10.1371/journal.ppat.1005937.
- Kirchdoerfer, R. N., Saphire, E. O., and Ward, A. B. (2019). Cryo-EM structure of the Ebola virus nucleoprotein – RNA complex research communications. 35, 340–347. doi:10.1107/S2053230X19004424.
- Kolakofsky, D. (2016). Paramyxovirus RNA synthesis , mRNA editing , and genome hexamer phase : A review. *Virology* 498, 94–98. doi:10.1016/j.virol.2016.08.018.
- Kolakofsky, D., Pelet, T., Garcin, D., Hausmann, S., Curran, J., and Roux, L. (1998). Paramyxovirus RNA synthesis and the requirement for hexamer genome length: the rule of six revisited. *J. Virol.* 72, 891–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9444980> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC124558>.
- Kolesnikova, L., Nanbo, A., Becker, S., and Kawaoka, Y. (2017). “Inside the Cell: Assembly of Filoviruses,” in *Current topics in microbiology and immunology*, 353–380. doi:10.1007/82\_2017\_15.
- Kondratowicz, A. S., Lennemann, N. J., Sinn, P. L., Davey, R. A., Hunt, C. L., Moller-Tank, S., et al. (2011). T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus. *Proc. Natl. Acad. Sci. U. S. A.* 108, 8426–31. doi:10.1073/pnas.1019030108.
- Kruse, T., Biedenkopf, N., Hertz, E. P. T., Dietzel, E., Stalman, G., López-Méndez, B., et al. (2018). The Ebola Virus Nucleoprotein Recruits the Host PP2A-B56 Phosphatase to Activate Transcriptional Support Activity of VP30. *Mol. Cell* 69, 136–145.e6. doi:10.1016/j.molcel.2017.11.034.
- Ksiazek, T. G., Rollin, P. E., Williams, A. J., Bressler, D. S., Martin, M. L., Swanepoel, R., et al. (1999). Clinical Virology of Ebola Hemorrhagic Fever (EHF): Virus, Virus Antigen, and IgG and IgM Antibody Findings among EHF Patients in Kikwit, Democratic Republic of the Congo, 1995. *J. Infect. Dis.* 179, S177–S187. doi:10.1086/514321.

- Kuhn, J. H., Adachi, T., Adhikari, N. K. J., Arribas, J. R., Bah, I. E., Bausch, D. G., et al. (2019a). New filovirus disease classification and nomenclature. *Nat. Rev. Microbiol.* 17, 261–263. doi:10.1038/s41579-019-0187-4.
- Kuhn, J. H., Amarasinghe, G. K., Basler, C. F., Bavari, S., Bukreyev, A., Chandran, K., et al. (2019b). ICTV virus taxonomy profile: Filoviridae. *J. Gen. Virol.* 100, 911–912. doi:10.1099/jgv.0.001252.
- Kurilla, M. G., Stone, H. O., and Keene, J. D. (1985). RNA sequence and transcriptional properties of the 3' end of the Newcastle disease virus genome. *Virology* 145, 203–212. doi:10.1016/0042-6822(85)90154-0.
- Le Mercier, P., Garcin, D., Garcia, E., and Kolakofsky, D. (2003). Competition between the Sendai virus N mRNA start site and the genome 3'-end promoter for viral RNA polymerase. *J. Virol.* 77, 9147–55. doi:10.1128/jvi.77.17.9147-9155.2003.
- Le Mercier, P., and Kolakofsky, D. (2019). Bipartite promoters and RNA editing of paramyxoviruses and filoviruses. *RNA* 25, 279–285. doi:10.1261/rna.068825.118.
- Lee, J. E., and Saphire, E. O. (2009). Ebolavirus glycoprotein structure and mechanism of entry. *Future Virol.* 4, 621–635. doi:10.2217/fvl.09.56.
- Leligdowicz, A., Fischer, W. A., Uyeki, T. M., Fletcher, T. E., Adhikari, N. K. J., Portella, G., et al. (2016). Ebola virus disease and critical illness. *Crit. Care* 20, 217. doi:10.1186/s13054-016-1325-2.
- Leppert, M., Rittenhouse, L., Perrault, J., Summers, D. F., and Kolakofsky, D. (1979). Plus and minus strand leader RNAs in negative strand virus-infected cells. *Cell* 18, 735–747. doi:10.1016/0092-8674(79)90127-2.
- Leroy, E. M., Gonzalez, J.-P., and Baize, S. (2011). Ebola and Marburg haemorrhagic fever viruses: major scientific advances, but a relatively minor public health threat for Africa. *Clin. Microbiol. Infect.* 17, 964–76. doi:10.1111/j.1469-0691.2011.03535.x.
- Leroy, E. M., Rouquet, P., Formenty, P., Souquière, S., Kilbourne, A., Froment, J.-M., et al. (2004). Multiple Ebola Virus Transmission Events and Rapid Decline of Central African Wildlife. *Science* (80-. ). 303, 387–390. doi:10.1126/science.1092528.
- Li, D., Willkomm, D. K., and Hartmann, R. K. (2009). Minor changes largely restore catalytic activity of archaeal RNase P RNA from *Methanothermobacter thermoautotrophicus*. *Nucleic Acids Res.* 37, 231–242. doi:10.1093/nar/gkn915.
- Li, T., and Pattnaik, A. K. (1999). Overlapping signals for transcription and replication at the 3' terminus of the vesicular stomatitis virus genome. *J. Virol.* 73, 444–52. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9847350> [Accessed August 29, 2019].
- Liang, B., Li, Z., Jenni, S., Rahmeh, A. A., Morin, B. M., Grant, T., et al. (2015). Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell*. doi:10.1016/j.cell.2015.06.018.
- Lier, C., Becker, S., and Biedenkopf, N. (2017). Dynamic phosphorylation of Ebola virus VP30 in NP-induced inclusion bodies. *Virology* 512, 39–47. doi:10.1016/j.virol.2017.09.006.
- Lo, T. Q., Marston, B. J., Dahl, B. A., and De Cock, K. M. (2017). Ebola: Anatomy of an Epidemic. *Annu. Rev. Med.* 68, 359–370. doi:10.1146/annurev-med-052915-015604.
- Louber, J., Kowalinski, E., Bloyet, L. M., Brunel, J., Cusack, S., and Gerlier, D. (2014). RIG-I self-oligomerization is either dispensable or very transient for signal transduction. *PLoS One* 9, e108770. doi:10.1371/journal.pone.0108770.

- Lytle, J. R., Wu, L., and Robertson, H. D. (2002). The Ribosome Binding Site of Hepatitis C Virus mRNA. *J. Virol.* 75, 7629–7636. doi:10.1128/jvi.75.16.7629-7636.2001.
- Madara, J. J., Han, Z., Ruthel, G., Freedman, B. D., and Harty, R. N. (2015). The multifunctional Ebola virus VP40 matrix protein is a promising therapeutic target. *Future Virol.* 10, 537–546. doi:10.2217/fvl.15.6.
- Malvy, D., McElroy, A. K., de Clerck, H., Günther, S., and van Griensven, J. (2019). Ebola virus disease. *Lancet* 393, 936–948. doi:10.1016/S0140-6736(18)33132-5.
- Marq, J.-B., Kolakofsky, D., and Garcin, D. (2010). Unpaired 5' ppp-nucleotides, as found in arenavirus double-stranded RNA panhandles, are not recognized by RIG-I. *J. Biol. Chem.* 285, 18208–16. doi:10.1074/jbc.M109.089425.
- Martin, B., Coutard, B., Guez, T., Paesen, G. C., Canard, B., Debart, F., et al. (2018). The methyltransferase domain of the Sudan ebolavirus L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. *Nucleic Acids Res.* 46, 7902–7912. doi:10.1093/nar/gky637.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10. doi:10.14806/ej.17.1.200.
- Martines, R. B., Ng, D. L., Greer, P. W., Rollin, P. E., and Zaki, S. R. (2015). Tissue and cellular tropism, pathology and pathogenesis of Ebola and Marburg viruses. *J. Pathol.* 235, 153–174. doi:10.1002/path.4456.
- Martinez, M. J., Volchkova, V. A., Raoul, H., Alazard-Dany, N., Reynard, O., and Volchkov, V. E. (2011). Role of VP30 Phosphorylation in the Ebola Virus Replication Cycle. *J. Infect. Dis.* 204, S934–S940. doi:10.1093/infdis/jir320.
- Martini, G. A. (1971). “Marburg Virus Disease. Clinical Syndrome,” in *Marburg Virus Disease* (Berlin, Heidelberg: Springer Berlin Heidelberg), 1–9. doi:10.1007/978-3-662-01593-3\_1.
- Matsumoto, Y., Ohta, K., Kolakofsky, D., and Nishio, M. (2017). A Point Mutation in the RNA-Binding Domain of Human Parainfluenza Virus Type 2 Nucleoprotein Elicits Abnormally Enhanced Polymerase Activity. *J. Virol.* 91, 1–14. doi:10.1128/jvi.02203-16.
- Matsumoto, Y., Ohta, K., Kolakofsky, D., and Nishio, M. (2018). The control of paramyxovirus genome hexamer length and mRNA editing. 461–467. doi:10.1261/rna.065243.117.a.
- Mbala-Kingebeni, P., Villabona-Arenas, C. J., Vidal, N., Likofata, J., Nsio-Mbeta, J., Makiala-Mandanda, S., et al. (2019). Rapid Confirmation of the Zaire Ebola Virus in the Outbreak of the Equateur Province in the Democratic Republic of Congo: Implications for Public Health Interventions. *Clin. Infect. Dis.* 68, 330–333. doi:10.1093/cid/ciy527.
- McElroy, A. K., Akondy, R. S., Davis, C. W., Ellebedy, A. H., Mehta, A. K., Kraft, C. S., et al. (2015). Human Ebola virus infection results in substantial immune activation. *Proc. Natl. Acad. Sci. U. S. A.* 112, 4719–4724. doi:10.1073/pnas.1502619112.
- McElroy, A. K., Mühlberger, E., and Muñoz-Fontela, C. (2018). Immune barriers of Ebola virus infection. *Curr. Opin. Virol.* 28, 152–160. doi:10.1016/j.coviro.2018.01.010.
- McGivern, D. R., Collins, P. L., and Fearn, R. (2005). Identification of Internal Sequences in the 3' Leader Region of Human Respiratory Syncytial Virus That Enhance Transcription and Confer Replication Processivity. *J. Virol.* 79, 2449–2460. doi:10.1128/jvi.79.4.2449-2460.2005.
- Mehedi, M., Falzarano, D., Seebach, J., Hu, X., Carpenter, M. S., Schnittler, H.-J., et al. (2011). A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J. Virol.* 85, 5406–14. doi:10.1128/JVI.02190-10.

- Mehedi, M., Hoenen, T., Robertson, S., Ricklefs, S., Dolan, M. A., Taylor, T., et al. (2013). Ebola Virus RNA Editing Depends on the Primary Editing Site Sequence and an Upstream Secondary Structure. *PLoS Pathog.* 9, e1003677. doi:10.1371/journal.ppat.1003677.
- Modrof, J., Becker, S., and Muhlberger, E. (2003). Ebola Virus Transcription Activator VP30 Is a Zinc-Binding Protein. *J. Virol.* 77, 3334–3338. doi:10.1128/jvi.77.5.3334-3338.2003.
- Modrof, J., Muhlberger, E., Klenk, H.-D., and Becker, S. (2002). Phosphorylation of VP30 impairs ebola virus transcription. *J. Biol. Chem.* 277, 33099–104. doi:10.1074/jbc.M203775200.
- Morin, B., Liang, B., Gardner, E., Ross, R. A., and Whelan, S. P. J. (2017). An In Vitro RNA Synthesis Assay for Rabies Virus Defines Ribonucleoprotein Interactions Critical for Polymerase Activity. *J. Virol.* 91. doi:10.1128/JVI.01508-16.
- Morin, B., Rahmeh, A. A., and Whelan, S. P. (2012). Mechanism of RNA synthesis initiation by the vesicular stomatitis virus polymerase. *EMBO J.* 31, 1320–1329. doi:10.1038/emboj.2011.483.
- Mühlberger, E. (2007). Filovirus replication and transcription. *Future Virol.* 2, 205–215. doi:10.2217/17460794.2.2.205.
- Mühlberger, E., Lötfering, B., Klenk, H. D., and Becker, S. (1998). Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. *J. Virol.* 72, 8756–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9765419> [Accessed August 27, 2019].
- Mühlberger, E., Trommer, S., Funke, C., Volchkov, V., Klenk, H.-D., and Becker, S. (1996). Termini of all mRNA species of Marburg virus: Sequence and secondary structure. *Virology* 223, 376–380. Available at: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed4&NEWS=N&AN=1996293636>.
- Mühlberger, E., Weik, M., Volchkov, V. E., Klenk, H. D., and Becker, S. (1999). Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J. Virol.* 73, 2333–42. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9971816> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC104478>.
- Murphy, S. K., Ito, Y., and Parks, G. D. (1998). A functional antigenomic promoter for the paramyxovirus simian virus 5 requires proper spacing between an essential internal segment and the 3' terminus. *J. Virol.* 72, 10–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9420195> [Accessed August 30, 2019].
- Murphy, S. K., and Parks, G. D. (1997). Genome Nucleotide Lengths That Are Divisible by Six Are Not Essential but Enhance Replication of Defective Interfering RNAs of the Paramyxovirus Simian Virus 5. *Virology* 232, 145–157. doi:10.1006/viro.1997.8530.
- Murphy, S. K., and Parks, G. D. (1999). RNA replication for the paramyxovirus simian virus 5 requires an internal repeated (CGNNNN) sequence motif. *J. Virol.* 73, 805–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9847393> [Accessed August 30, 2019].
- Nakkazi, E. (2018). Randomised controlled trial begins for Ebola therapeutics. *Lancet* 392, 2338. doi:10.1016/S0140-6736(18)33011-3.
- Negredo, A., Palacios, G., Vázquez-Morón, S., González, F., Dopazo, H., Molero, F., et al. (2011). Discovery of an Ebolavirus-Like Filovirus in Europe. *PLoS Pathog.* 7, e1002304. doi:10.1371/journal.ppat.1002304.



- Neumann, G., Geisbert, T. W., Ebihara, H., Geisbert, J. B., Daddario-DiCaprio, K. M., Feldmann, H., et al. (2007). Proteolytic processing of the Ebola virus glycoprotein is not critical for Ebola virus replication in nonhuman primates. *J. Virol.* 81, 2995–8. doi:10.1128/JVI.02486-06.
- Neumann, G., Watanabe, S., and Kawaoka, Y. (2009). Characterization of Ebolavirus regulatory genomic regions. *Virus Res.* 144, 1–7. doi:10.1016/j.virusres.2009.02.005.
- Normile, D. (2009). Emerging infectious diseases. Scientists puzzle over Ebola-Reston virus in pigs. *Science* 323, 451. doi:10.1126/science.323.5913.451a.
- Noton, S. L., and Fearn, R. (2015). Initiation and regulation of paramyxovirus transcription and replication. *Virology* 479–480, 545–554. doi:10.1016/j.virol.2015.01.014.
- Ogino, T., and Green, T. J. (2019). RNA Synthesis and Capping by Non-segmented Negative Strand RNA Viral Polymerases: Lessons From a Prototypic Virus. *Front. Microbiol.* 10, 1490. doi:10.3389/fmicb.2019.01490.
- Oh, S., Onomoto, K., Wakimoto, M., and Onoguchi, K. (2016). Leader-Containing Uncapped Viral Transcript Activates RIG-I in Antiviral Stress Granules. 1–22. doi:10.1371/journal.ppat.1005444.
- Olsen, M. E., Cressey, T. N., Mühlberger, E., and Connor, J. H. (2018). Differential Mechanisms for the Involvement of Polyamines and Hypusinated eIF5A in Ebola Virus Gene Expression. *J. Virol.* 92, e01260-18. doi:10.1128/JVI.01260-18.
- Patton, J. T., Davis, N. L., and Wertz, G. W. (1984). N protein alone satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. *J. Virol.* 49, 303–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/6319730> [Accessed September 4, 2019].
- Pelet, T., Delenda, C., Gubbay, O., Garcin, D., and Kolakofsky, D. (1996). Partial characterization of a Sendai virus replication promoter and the role of six. *Virology* 224, 405–14. doi:10.1006/viro.1996.0547.
- Perez, J. T., Varble, A., Sachidanandam, R., Zlatev, I., Manoharan, M., García-Sastre, A., et al. (2010). Influenza A virus-generated small RNAs regulate the switch from transcription to replication. *Proc. Natl. Acad. Sci. U. S. A.* 107, 11525–30. doi:10.1073/pnas.1001984107.
- Perrault, J., Clinton, G. M., and McClure, M. A. (1983). RNP template of vesicular stomatitis virus regulates transcription and replication functions. *Cell* 35, 175–185. doi:10.1016/0092-8674(83)90220-9.
- Peters, D., Müller, G., and Slenczka, W. (1971). “Morphology, Development, and Classification of the Marburg Virus,” in *Marburg Virus Disease* (Berlin, Heidelberg: Springer Berlin Heidelberg), 68–83. doi:10.1007/978-3-662-01593-3\_10.
- Pigott, D. C. (2005). Hemorrhagic Fever Viruses. *Crit. Care Clin.* 21, 765–783. doi:10.1016/j.ccc.2005.06.007.
- Plumet, S., Herschke, F., Bourhis, J. M., Valentin, H., Longhi, S., and Gerlier, D. (2007). Cytosolic 5'-triphosphate ended viral leader transcript of measles virus as activator of the RIG I-mediated interferon response. *PLoS One* 2, 1–11. doi:10.1371/journal.pone.0000279.
- Ramakers, C., Ruijter, J. M., Lekanne Deprez, R. H., and Moorman, A. F. M. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* 339, 62–66. doi:10.1016/S0304-3940(02)01423-4.
- Reed, D. S., Hensley, L. E., Geisbert, J. B., Jahrling, P. B., and Geisbert, T. W. (2004). Depletion of Peripheral Blood T Lymphocytes and NK Cells During the Course of

- Ebola Hemorrhagic Fever in Cynomolgus Macaques. *Viral Immunol.* 17, 390–400. doi:10.1089/vim.2004.17.390.
- Rhein, B. A., and Maury, W. J. (2015). Ebola Virus Entry into Host Cells: Identifying Therapeutic Strategies. *Curr. Clin. Microbiol. Reports* 2, 115–124. doi:10.1007/s40588-015-0021-3.
- Rougeron, V., Feldmann, H., Grard, G., Becker, S., and Leroy, E. M. (2015). Ebola and Marburg haemorrhagic fever. *J. Clin. Virol.* 64, 111–119. doi:10.1016/j.jcv.2015.01.014.
- Sadewasser, A., Dietzel, E., Michel, S., Klüver, M., Helfer, M., Thelemann, T., et al. (2019). Anti-Niemann Pick C1 Single-Stranded Oligonucleotides with Locked Nucleic Acids Potently Reduce Ebola Virus Infection In Vitro. *Mol. Ther. Nucleic Acids* 16, 686–697. doi:10.1016/j.omtn.2019.04.018.
- Sanchez, A., and Kiley, M. P. (1987). Identification and analysis of ebola virus messenger RNA. *Virology* 157, 414–420. doi:10.1016/0042-6822(87)90283-2.
- Sanchez, A., Kiley, M. P., Holloway, B. P., and Auperin, D. D. (1993). Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res.* 29, 215–240. doi:10.1016/0168-1702(93)90063-S.
- Schlee, M., Roth, A., Hornung, V., Hagmann, C. A., Wimmenauer, V., Barchet, W., et al. (2009). Recognition of 5' Triphosphate by RIG-I Helicase Requires Short Blunt Double-Stranded RNA as Contained in Panhandle of Negative-Strand Virus. *Immunity* 31, 25–34. doi:10.1016/j.immuni.2009.05.008.
- Schlereth, J. (2015). Charakterisierung der RNA-Bindungsaktivität von VP30 im Hinblick auf die virale Transkriptionsregulation des Ebolavirus. *Dissertation*.
- Schlereth, J., Grünweller, A., Biedenkopf, N., Becker, S., and Hartmann, R. K. (2016). RNA binding specificity of Ebola virus transcription factor VP30. *RNA Biol.* 13, 783–798. doi:10.1080/15476286.2016.1194160.
- Schmidt, A., Schwerd, T., Hamm, W., Hellmuth, J. C., Cui, S., Wenzel, M., et al. (2009). 5'-triphosphate RNA requires base-paired structures to activate antiviral signaling via RIG-I. *Proc. Natl. Acad. Sci.* 106. doi:10.1073/pnas.0900971106.
- Shabman, R. S., Hoenen, T., Groseth, A., Jabado, O., Binning, J. M., Amarasinghe, G. K., et al. (2013). An Upstream Open Reading Frame Modulates Ebola Virus Polymerase Translation and Virus Replication. *PLoS Pathog.* 9. doi:10.1371/journal.ppat.1003147.
- Shabman, R. S., Jabado, O. J., Mire, C. E., Stockwell, T. B., Edwards, M., Mahajan, M., et al. (2014). Deep Sequencing Identifies Noncanonical Editing of Ebola and Marburg Virus RNAs in Infected Cells. 5, 1–11. doi:10.1128/mBio.02011-14.Editor.
- Shi, M., Lin, X.-D., Chen, X., Tian, J.-H., Chen, L.-J., Li, K., et al. (2018). The evolutionary history of vertebrate RNA viruses. *Nature* 556, 197–202. doi:10.1038/s41586-018-0012-7.
- Shu, T., Gan, T., Bai, P., Wang, X., Qian, Q., Zhou, H., et al. (2019). Ebola virus VP35 has novel NTPase and helicase-like activities. *Nucleic Acids Res.* 47, 5837–5851. doi:10.1093/nar/gkz340.
- Sivapalasingam, S., Kamal, M., Slim, R., Hosain, R., Shao, W., Stoltz, R., et al. (2018). Safety, pharmacokinetics, and immunogenicity of a co-formulated cocktail of three human monoclonal antibodies targeting Ebola virus glycoprotein in healthy adults: a randomised, first-in-human phase 1 study. *Lancet Infect. Dis.* 18, 884–893. doi:10.1016/S1473-3099(18)30397-9.
- Suder, E., Furuyama, W., Feldmann, H., Marzi, A., and de Wit, E. (2018). The vesicular

- stomatitis virus-based Ebola virus vaccine: From concept to clinical trials. *Hum. Vaccin. Immunother.* 14, 2107–2113. doi:10.1080/21645515.2018.1473698.
- Sugita, Y., Matsunami, H., Kawaoka, Y., Noda, T., and Wolf, M. (2018). Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex at 3.6 Å resolution. *Nature*. doi:10.1038/s41586-018-0630-0.
- Sztuba-Solinska, J., Diaz, L., Kumar, M. R., Kolb, G., Wiley, M. R., Jozwick, L., et al. (2016). A small stem-loop structure of the Ebola virus trailer is essential for replication and interacts with heat-shock protein A8. *Nucleic Acids Res.* 44, 9831–9846. doi:10.1093/nar/gkw825.
- Takahashi, K., Halfmann, P., Oyama, M., Kozuka-Hata, H., Noda, T., and Kawaoka, Y. (2013). DNA Topoisomerase 1 Facilitates the Transcription and Replication of the Ebola Virus Genome. *J. Virol.* 87, 8862–8869. doi:10.1128/jvi.03544-12.
- Takamatsu, Y., Kolesnikova, L., and Becker, S. (2018). Ebola virus proteins NP, VP35, and VP24 are essential and sufficient to mediate nucleocapsid transport. *Proc. Natl. Acad. Sci. U. S. A.* 115, 1075–1080. doi:10.1073/pnas.1712263115.
- Takamatsu, Y., Kolesnikova, L., Halwe, S., Baumeister, S., Noda, T., Biedenkopf, N., et al. (2019). Serine-arginine protein kinase 1 regulates Ebola virus transcription. *bioRxiv*, 717371. doi:10.1101/717371.
- Tapparel, C., Maurice, D., and Roux, L. (1998). The activity of Sendai virus genomic and antigenomic promoters requires a second element past the leader template regions: a motif (GNNNNN)<sub>3</sub> is essential for replication. *J. Virol.* 72, 3117–28. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9525637> [Accessed August 28, 2019].
- Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Am. Math. Soc. Lect. Math. Life Sci.* 17, 57–86. doi:citeulike-article-id:4801403.
- Tawar, R. G., Duquerroy, S., Vonnrhein, C., Varela, P. F., Damier-Piolle, L., Castagné, N., et al. (2009). Crystal Structure of a Nucleocapsid-Like Nucleoprotein-RNA Complex of Respiratory Syncytial Virus. *Science* (80-. ). 326, 1279–1283. doi:10.1126/science.1177634.
- Tchesnokov, E. P., Raeisimakiani, P., Ngure, M., Marchant, D., and Götte, M. (2018). Recombinant RNA-Dependent RNA Polymerase Complex of Ebola Virus. 1–9. doi:10.1038/s41598-018-22328-3.
- Tekes, G., Rahmeh, A. A., and Whelan, S. P. J. (2011). A Freeze Frame View of Vesicular Stomatitis Virus Transcription Defines a Minimal Length of RNA for 5' Processing. *PLoS Pathog.* 7, e1002073. doi:10.1371/journal.ppat.1002073.
- The R Development Core Team (2011). R: A Language and Environment for Statistical Computing Reference Index. R Foundation for Statistical Computing Available at: <https://stat.ethz.ch/pipermail/r-help/2008-May/161481.html> [Accessed November 29, 2019].
- Tomescu, A. I., Robb, N. C., Hengrung, N., Fodor, E., and Kapanidis, A. N. (2014). Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc. Natl. Acad. Sci.* 111, E3335–E3342. doi:10.1073/pnas.1406056111.
- Towner, J. S., Rollin, P. E., Bausch, D. G., Sanchez, A., Crary, S. M., Vincent, M., et al. (2004). Rapid Diagnosis of Ebola Hemorrhagic Fever by Reverse Transcription-PCR in an Outbreak Setting and Assessment of Patient Viral Load as a Predictor of Outcome. *J. Virol.* 78, 4330–4341. doi:10.1128/jvi.78.8.4330-4341.2004.
- Tremaglio, C. Z., Noton, S. L., Deflube, L. R., and Fearn, R. (2013). Respiratory Syncytial Virus Polymerase Can Initiate Transcription from Position 3 of the Leader

- Promoter. *J. Virol.* 87, 3196–3207. doi:10.1128/JVI.02862-12.
- Umbach, J. L., Yen, H.-L., Poon, L. L. M., and Cullen, B. R. (2010). Influenza A virus expresses high levels of an unusual class of small viral leader RNAs in infected cells. *MBio* 1. doi:10.1128/mBio.00204-10.
- Vidal, S., and Kolakofsky, D. (1989). Modified model for the switch from Sendai virus transcription to replication. *J. Virol.* 63, 1951–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2539496> [Accessed August 30, 2019].
- Volchkov, V. E., Feldmann, H., Volchkova, V. A., and Klenk, H. D. (1998). Processing of the Ebola virus glycoprotein by the proprotein convertase furin. *Proc. Natl. Acad. Sci. U. S. A.* 95, 5762–7. doi:10.1073/pnas.95.10.5762.
- Volchkov, V. E., Volchkova, V. A., Chepurinov, A. A., Blinov, V. M., Dolnik, O., Netesov, S. V., et al. (1999). Characterization of the L gene and 5' trailer region of Ebola virus. *J. Gen. Virol.* 80, 355–362. doi:10.1099/0022-1317-80-2-355.
- Vulli  moz, D., Cordey, S., Mottet-Osman, G., and Roux, L. (2005). Nature of a paramyxovirus replication promoter influences a nearby transcription signal. *J. Gen. Virol.* 86, 171–180. doi:10.1099/vir.0.80435-0.
- Vulli  moz, D., and Roux, L. (2001). ‘Rule of six’: how does the Sendai virus RNA polymerase keep count? *J. Virol.* 75, 4506–18. doi:10.1128/JVI.75.10.4506-4518.2001.
- Wan, W., Kolesnikova, L., Clarke, M., Koehler, A., Noda, T., Becker, S., et al. (2017). Structure and assembly of the Ebola virus nucleocapsid. *Nature* 551, 394–397. doi:10.1038/nature24490.
- Watt, A., Moukambi, F., Banadyga, L., Groseth, A., Callison, J., Herwig, A., et al. (2014). A Novel Life Cycle Modeling System for Ebola Virus Shows a Genome Length-Dependent Role of VP24 in Virus Infectivity. *J. Virol.* 88, 10511–10524. doi:10.1128/JVI.01272-14.
- Weik, M., Enterlein, S., Schlenz, K., and M  hlberger, E. (2005). The Ebola Virus Genomic Replication Promoter Is Bipartite and Follows the Rule of Six. *J. Virol.* 79, 10660–10671. doi:10.1128/jvi.79.16.10660-10671.2005.
- Weik, M., Modrof, J., Klenk, H.-D., Becker, S., and M  hlberger, E. (2002). Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J. Virol.* 76, 8532–9. doi:10.1128/JVI.76.17.8532.
- West, T. E., and von Saint Andr  -von Arnim, A. (2014). Clinical Presentation and Management of Severe Ebola Virus Disease. *Ann. Am. Thorac. Soc.* 11, 1341–1350. doi:10.1513/AnnalsATS.201410-481PS.
- Whelan, S. P. J., Barr, J. N., and Wertz, G. W. (2004). Transcription and replication of nonsegmented negative-strand RNA viruses. *Curr. Top. Microbiol. Immunol.* 283, 61–119. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15298168>.
- Whelan, S. P., and Wertz, G. W. (1999). Regulation of RNA synthesis by the genomic termini of vesicular stomatitis virus: identification of distinct sequences essential for transcription but not replication. *J. Virol.* 73, 297–306. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9847333> [Accessed September 5, 2019].
- World Health Organization (2019). Ebola Virus disease: Democratic Republic of the Congo. 1–9. Available at: [https://apps.who.int/iris/bitstream/handle/10665/324996/SITREP\\_EVD\\_DRC\\_20190528-eng.pdf?ua=1%0Ahttps://www.afro.who.int/health-topics/ebola-virus-disease](https://apps.who.int/iris/bitstream/handle/10665/324996/SITREP_EVD_DRC_20190528-eng.pdf?ua=1%0Ahttps://www.afro.who.int/health-topics/ebola-virus-disease).
- Xu, W., Luthra, P., Wu, C., Batra, J., Leung, D. W., Basler, C. F., et al. (2017). Ebola virus VP30 and nucleoprotein interactions modulate viral RNA synthesis. *Nat. Commun.* 8, 1–11. doi:10.1038/ncomms15576.

- Yang, D., and Leibowitz, J. L. (2015). The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res.* doi:10.1016/j.virusres.2015.02.025.
- Yang, X.-L., Tan, C. W., Anderson, D. E., Jiang, R.-D., Li, B., Zhang, W., et al. (2019). Characterization of a filovirus (Měnglà virus) from Rousettus bats in China. *Nat. Microbiol.* 4, 390–395. doi:10.1038/s41564-018-0328-y.
- Zhu, F.-C., Wurie, A. H., Hou, L.-H., Liang, Q., Li, Y.-H., Russell, J. B. W., et al. (2017). Safety and immunogenicity of a recombinant adenovirus type-5 vector-based Ebola vaccine in healthy adults in Sierra Leone: a single-centre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* 389, 621–628. doi:10.1016/S0140-6736(16)32617-4.

## 8 ANHANG

### 8.1 Ergänzende Klonierungsstrategien und Primer

In der nachfolgenden Tabelle sind Klonierungsstrategien und Primer für die Herstellung der bicistronischen Konstrukte bici\_NP-GP und bici\_NP-VP24 aus Abb. 20 (Kapitel 6, Ausblick) angegeben. Alle weiteren Klonierungsstrategien sowie verwendete Primer sind den Ergänzungen der Manuskripte zu entnehmen. Das EBOV bicistronische RC-Wt-Minigenom pA-3E5E-RLuc-FLuc (Bach et al., 2019c/Manuskript 3, Kapitel 4.3) wird abgekürzt als bici\_NP-VP35 bezeichnet. Dickgedruckte, unterstrichene Nukleotide zeigen Insertionen an. Im Falle der Restriktionsklonierung verweisen dick- und kursivdruckte Nt auf Restriktionssequenzen.

Tab. 39: Ergänzende Klonierungsstrategien und Primer.

Nr.	Name	Klonierungsstrategie	Primer fwd	Primer fwd (5' to 3')	Primer rev	Primer rev (5' to 3')
Restriktionsklonierung						
1	bici_NP-GP	Templat-Rückgrat: bici_NP-VP35; Templat-Insert: GeneArt Strings DNA fragment_GP	GeneArt Strings DNA fragment_GP (Thermo Fisher Scientific; TFS): TTC <b>CTCGAG</b> CTTGTTAGACCAGCTTTTCTTTTACATTGAAGACAATAAGCT TCGAGTGCTAGAGAATTCAATTGACGTAAGGATGATGAAGATTACGCTCA CTGTCGGCTTAATCTTCATCATTAGTTTTTCTTAATTATAAACGATCGTG TAAATGTTGGGCAGGCTAGTAGGTAAGTTATAAAAGCATAATGTAATTTAA TTTAGATAAATCGCTAGTTAGAATATTGAGTTACAATTATGGCTATAACTG TGGCTTAACGCCTATTTGCATTAAGACTCAAATTGAACCTTTGAAAAACGGC TAGTAATAATAAGCTAGCTTTAAATTCAAAAATAAGAATCACTTTTATTGTG ATAATAATTCACACCTTAGACATCAAAAATTCCTCTGTTTCGTTTCCTTGA CTACTTAATGTTAGCTATTTGTCGGTTGAATCATCCCATTGTTCCATGCTC AT <b>GCGGCCGCGAG</b>			
Restriktionsverdau mit <i>Xho</i> I (5'- CTCGAG -3') und <i>Not</i> I (5'- GCGGCCGC -3'), dann Ligation mit T4 DNA ligase.						
Gibson Assembly mit 2x Gibson Assembly® Master Mix (NEB, #E2611S) nach Herstelleranleitung.						
2	bici_NP-VP24	1) PCR mit p3E5E_VP24 (Bach et al., 2019a); 2) PCR mit bici_NP-VP35  unter Verwendung der Phusion DNA-Polymerase (TFS)	2) pS_0153 2) pS_0155	1) AAT TCA ATT GAC GTA AGG <b><u>ATG</u></b> <b><u>ATG AAG AAT GCC</u></b> <b><u>TCT TGG GGC</u></b>  2) <b><u>CCT CCG CAT</u></b> <b><u>TAA TCT TCA TCA</u></b> TTA GTT TTT TCT TAA TTA TAA AAC GAT CGT GTA AAT G	1) pS_0154 2) pS_0156	1) ATA ATT AAG AAA AAA CTA <b><u>ATG</u></b> <b><u>ATG AAG ATT AAT</u></b> <b><u>GCG GAG GTC TG</u></b>  2) <b><u>CCC AAG AGG</u></b> <b><u>CAT TCT TCA TCA</u></b> TCC TTA CGT CAA TTG AAT TCT CTA GC
Austausch des VP35-HPs am 2. Cistron mit dem VP24-HP			Sequenz VP24: GATGAAGAATGCCTCTTGGGGCCTAATCTGAATAATAAGGTTTATTCTTAT CAGACCTCCGCATTAATCTTCATC			

## 8.2 Plasmidsequenz pA-3E5E-GLuc-GFP

Das Plasmid pA-3E5E-GLuc-GFP diente neben dem pCAGGS\_Luc2-Plasmid (Biedenkopf and Hoenen, 2017) als Templat für die Klonierung des bicistronischen RC-Wt-Minigenoms pA-3E5E-RLuc-FLuc (Bach et al., 2019c/Manuskript 3, Kapitel 4.3) .

Quelle	Sequenz (5' to 3')
Dr. Nadine Biedenkopf, AG Becker, Institut für Virologie Marburg	<p>ATCGATCCCGGGTTAATACGACTCACTATAGTGGACACACAAAAAGAAGAAATAGATTTATTTTTAAA  TTTTTGTGTGCGACCATTTTTCCAGGAATCCTTTTTGCAACGTTTATTCTGGCGAGCCTTCTGATGAGC  GTGGTCAATGTCTTAATTATCATTAAACACGAAGATTATTCTAGTCTGCTATATTGTAAAGAAGGTATT  TCTGATTTTACAGTCCTGCCTTTTCTTTTAATTTTATCATTCTGTATGAAATAGTAGCTTACTAGACAAAA  ATGTTAAGGTTATTACACCTAATCTCTTATTTTACACAAAGATTAAGGCTATCACCGTTTTTAAACGAC  AAATTAAAGTAATGATAATTATAAAAAAATTAATCTTAAAGATCAATTTTGTAAAGAGTTGTGTTATC  AACCAAAACACTATTCCATCTGACAGGATATTGATACAACATTATTATCAATAACAATATGAGCCCAG  ACCTTTTCGTTAAAGCTTAATTATAAAGAGTGCAGAGTTTATTATGTTGCGTTAAATTCATTGCGGCAGTA  ATATGTCACACATGCTGCATTGTGAATTTTTTATTATCACGCAGGGAGAGAGGCTAAATATAGTATAT  AATATAGGGAAGTCTTTTCGTTATTTAAATAGTTGTTTATTGAGATTAGATCAATATGATGTATGAGAGCA  ATTTATGAGTTTTTTATAATCTGTATGTTAATAACCAACCTTTGCAAGTATCAGGATACTATGCGGCCGC  TCACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAAGTCCAGCAGGACCAT  GTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGTTGTGCGGG  CAGCAGCACGGGGCGGCTCGCGGATGGGGTGTCTGCTGGTAGTGGTGGCGGAGCTGCACGCTGC  CGTCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTGCGCCATGATA  TAGACGTTGTGGCTGTTGTAGTTGACTCCAGCTTGCCCCAGGATGTTGCCGCTCCTCCTTGAAGTC  GATGCCCTTCAGCTCGATGCGGTTACACAGGGGTGTCGCCCTCGAAGTTCACCTCGGCGCGGGTCTTG  TAGTTGCCGTCGTCCTTGAAGAAGATGGTGCCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGA  AGAAGTCGTGCTGCTCATGTGGTGGGGTACGGGCTGAAGCACTGCACGCCGTAGGTACAGGGTGG  TCACGAGGGTGGGCCAGGGCACGGGCAGCTTGGCGGTGGTGCAGATGAAGTTCAGGGTCAGCTTGC  CGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAAGTGTGGCCGTTTACGTCGCCGTCCA  GCTCGACAGGATGGGCACACCCCGGTGAACAGCTCCTCGCCCTTGTCAACATCTCGAGCATCTT  GTTAGACAGCTTTTCTTTTACATTGAAGACAATAAGCTTCGAGTGCTAGAGAATTCAATGACGTAAG  GATGATGAAGTTTTAATCTTCATCATTAGTTTTTCTTAATTATAAACGATCGTGTAATGTTGGGCA  GGCTAGTAGGTAAGTTATAAAAGCATAATGTAATTTAATTAGATAAATCGCTAGTTAGAATATTGAGTT  ACAATTATGGCTATAACTGTGGCTTAACGCCTATTTGCATTAAGACTCAAATGAACTTTGAAAAACGG  CTAGATAATAAGCTAGCTTTAAATTCAAAAATAAGAAATCACTTTTATTGTAGATAAATACACTTTA  GACATCAAAAAATCTTCTGTTTTCGTTTCTTCTGACTACTTAATGTTAGCTATTTGTCGGTTGAATCATCC  CATTGTTCCATGCTCATGCGGCCGCTTAGTCACCACCGGCCCTTATCTTGTCCACCTGGCCCTG  GATCTTGTGTCGCAAGGTGCGACAGCGTTGCGGCAGCCACTTCTTGAGCAGGTGAGAACACTGCACG  TTGAGAACCCCTTTGAGGCAGCCAGTTGTGAGTCCACACACAGATCGACCTGTGCGATGACGTGCT  CCATGGGCTCCAAGTCTTGAACCCAGGAATCTCAGGAATGTCGACGATCGCCTCGCCTATGCCGCC  CTGTGCGGACTCTTGTGCGCTTCGTAGGTGTGGCAGCGTCTGGGATGAAGTCTTTCATCTTGGGC  GTGCACTTGATGTGGGACAGGCAGATCAGACAGCCCTGGTGACGCCAGCTTTCGGGGCATTGGCTT  CCATCTCTTTGAGCACCTCCAGCGGACGCTTCTTCCCGGGCAACTTCCCGGCTGAGCATCGAGATC  CGTGTGCGGAAGTTGCTGGCCACGGCCAGCATGTTGAAGTCTTCGTTGCTCGGTGCTTGGCC  TCGGCCACAGCGATGCAGATCAGGGCAAACAGAAGTTGACTCCCATCATATGGGTGGTATGAGACA  TACTCGGAATTTGTGATTCCGAGCAATTTGTCTTTAACTATTTCCAAATGTTTTACCCCAAGCTTTAGG  GTTGTTGAATCTTCAATCCCCTAAGTGTTCAGTTCAAAGTCAAACCAAGGTTAATTTCAATCTTGTGTTG  ACTGTGAAGTAAATGCTGCTCTTAAACATAATGTTTCTATTAAAGTTGTTGGATCACTTGTGATGCTG  GAGGTGGAAGTGTGATGTGATAAGACCTAGGTGTTGACAAGGGATTTTCAACTGAGCACACTGATGC  CCTTGCTCCCTTCGGAGACCTAGGTTGTTCTCTATCTTTGTGATGTGGCTCTGAAACAAACAGGTGT  GATTACAGTAACAATTTCAATTTAAATTCGATATAAATTTCAATGAGAGGAAATTTAATCTTCTCTCA  TAGTTATTGCGACACAAAAGATCCTAAAAATCTCTTTCTTTTGTGTGTCGGGGTCCGGATGGCAT  CTCCACCTCCTCGCGTCCGACCTGGGCATCCGAAGGAGGACGTCGTCCACTCGGATGGCTAAGGG  AGAGCTCAAAAAAAGGATCCGGCTGTAAACAAAGCCGAAAGGAAGCTGAGTTGGCTGTGCCACC  GCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTGTGTAAG  GAGGAATAATATCCGGATCGAGATCCTCTAGATCGACCTGCAGGCATGCAAGCTGTATTACATAGT  GTCACCTAAATCGTATGTATGATACATAAGGTTATGTATTAATTGAGCCGCTTCTAACGCAATA  TGTACAAGCCTAATTGTGTAGCATCTGGCTTACTGAAGCAGACCCTATCATCTCTCTGTAAGTACC  GTCAGAGTCGGTTTGGTTGGACGAACCTTCTGAGTTTCTGGTAACGCCGTTCCGCACCCCGGAAATG  GTCAGCGAACCAATCAGCAGGGTCATGCTGAGCCAGATCCTCTACGCCGACGCATCGTGGCCGGC  ATCACCGGCGCCACAGGTGCGGTTGCTGGCGCTATATCGCCGACATCACCGATGGGGAAGATCGG  GCTCGCCACTTCCGGCTCATGAGCGCTTGTTCGGCGTGGGTATGGTGGCAGGCCCGTGGCCGGG  GGACTGTTGGGCGCCATCTCCTTGACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACT  ACTGGGCTGCTTCTAATGCAGGAGTCGATAAGGGAGAGCGTCGATATGGTGCATCTCAGTACAA  TCTGCTGTATGCCGCATAGTTAAGCCAGCCCGACACCCGCAACACCCGCTGAGCTGGGGAAGTAC  GGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCA  GAGGTTTTACCGTCATCACCGAAACGCGGAGACGAAAGGGCCTCGTGATACGCCTATTTTATAG  GTTAATGTCATGATAAATAGGTTTCTTAGACGTGAGGTGGCACTTTTCCGGGAAATGTGCGCGGAAC  CCCTATTTGTTATTTTCTAAATACATTCAAATATGATCCGCTCATGAGACAAATACCTGATTAATG  CTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATCCCTTTTTTG  CGGCATTTTGCCTTCTGTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAG</p>

	<p> TTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC  CCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTG  ACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACC  AGTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAAACCATGA  GTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTG  CACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAA  ACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAAC TGCGGA  ACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCAC  TTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC  TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACG  GGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGC  ATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAA  AGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCAC  TGAGCGTCAGACCCCGTAGAAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTG  CTGCTTGCAACAAAAAACCACCGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAACTC  TTTTCCGAAGGTAAC TGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAG  TTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGT  GGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAG  GCGCAGCGGTGCGGGCTGAACGGGGGGTTCTGTGCACACAGCCCAGCTTGGAGCGAACGACCTACAC  CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGGAGAAAGGCGGA  CAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACG  CCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCG  TCAGGGGGGGCGGAGCCTATGGAACACGCCAGCAACGCGGCCTTTTACGGTTCCTGGCCTTTTGCT  GGCCTTTTGCTCACATGTTCTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTG  AGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGAGCGAGTCAGTGAGCGAGGAAGCG  GAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCGTTGCCGATTCAATTAATGCAGGGGGAT </p>
--	--



## 8.3 Manuskripte und Veröffentlichungen

### 8.3.1 Manuskripte

- a) **Bach, S.**, Biedenkopf, N., Grünweller, A., Becker, S., Hartmann, R. K. (2019). Hexamer phasing governs transcription initiation in the 3'-leader of Ebola virus RNA. Eingereicht am 22.10.2019, akzeptiert zur Publikation am 08.01.2020.
- b) **Bach, S.**, Biedenkopf, N., Becker, S., Hartmann, R. K. Secondary structure formation influences EBOV transcription initiation and replication in a length and stability-dependent manner. In Vorbereitung.
- c) **Bach, S.**, Schlereth, J., Biedenkopf, N., Becker, S., Lechner, M., Hartmann, R. K. Ebola virus transcriptome analysis reveals insights into transcriptional regulation. In Vorbereitung.
- d) **Bach, S.**, Biedenkopf, N., Becker, S., Hartmann, R. K. Sequence, structure and length constraints in the spacer region of the genomic EBOV promoter on VP30 dependency of transcription and the balance of transcription and replication. In Vorbereitung.

### 8.3.2 Veröffentlichungen

- Damm, K., **Bach, S.**, Müller, K. M. H., Klug, G., Burenina, O. Y., Kubareva, E. A., et al. (2015a). Impact of RNA isolation protocols on RNA detection by northern blotting. doi:10.1007/978-1-4939-2547-6\_4.
- Damm, K., **Bach, S.**, Müller, K. M. H., Klug, G., Burenina, O. Y., Kubareva, E. A., et al. (2015b). Improved northern blot detection of small RNAs using EDC crosslinking and DNA/LNA probes. doi:10.1007/978-1-4939-2547-6\_5.

## 8.4 Konferenzbeiträge

### Vorträge und Posterbeiträge

- 10/2018 **Poster**; Stein, J.S., Bach, S.; Becker, S., Hartmann, R.K., *VP30 dependency in Ebola virus transcription*, RNA biochemistry meeting 2018, Bonn, Deutschland
- 09/2018 **Poster**; Bach, S., Lechner, M., Grünweller, A., Becker, S., Biedenkopf, N., Hartmann, R. K., *Interplay of VP30 and RNA secondary structures at the transcription start sites of the Ebola virus genome*, International Symposium on RNA viruses at the host interface, Marburg, Deutschland
- 06/2018 **Vortrag**; Bach, S., Grünweller, A., Becker, S., Biedenkopf, N., Hartmann, R. K., *Advanced analyses on the role of VP30 and RNA structures in EBOV transcription*, NSV 2018, Verona, Italien (gefördert durch DAAD und MID)
- 09/2017 **Vortrag**; Bach, S., Grünweller, A., Becker, S., Biedenkopf, N., Hartmann, R. K., *Interplay of VP30 and RNA secondary structures at the transcription start sites of the ZEBOV genome*, 9<sup>th</sup> International Symposium on Filoviruses, Marburg, Deutschland
- 09/2017 **Poster**; Bach, S., Pilgram J., Lange-Grünweller, K., Lechner, M., Schulte, L. N., Grünweller, A., Hartmann, R. K., *Targeting human pathogenic ncRNAs with membrane-permeable palmitoylated 2'-Amino-LNA Gapmers*, Medical RNomics Symposium 2017, Gießen, Deutschland
- 06/2017 **Vortrag**; Bach, S., Grünweller, A., Hartmann, R. K., *Was RNA alles kann*, Karl-von-Frisch Preisverleihung, Marburg, Deutschland
- 09/2016 **Vortrag**; Bach, S., Schulte, L. N., Grünweller, A., Hartmann, R. K., *Antisense strategies against ncRNAs*. Summer School Regenerative Medicine, Shanghai, China (gefördert durch DAAD)

## **8.5 Lebenslauf**

Diese Seite enthält persönliche Daten. Sie ist deshalb nicht Bestandteil der Online-Veröffentlichung

---

## 8.6 Eidesstattliche Erklärung

Ich versichere, dass ich meine Dissertation

„Analyse des Ebolavirus 3'-Leader-Promotors – mechanistische Beiträge von VP30, Promotorarchitektur, RNA-Sequenzelementen und RNA-Sekundärstrukturen zur viralen Transkriptions- und Replikationsinitiation“

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 17.01.2020

.....  
(Simone Bach)

## 8.7 Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen, mich bei all jenen zu bedanken, die mich durch meine Promotionszeit begleitet und geleitet haben.

Zuallererst möchte ich mich bei meinem Doktorvater, Prof. Dr. Roland K. Hartmann, bedanken. Lieber Roland, danke, dass ich meine Doktorarbeit in deiner Arbeitsgruppe anfertigen durfte. Vielen, vielen Dank für deine Unterstützung, deine wertvollen Anregungen und, dass du dir stets sehr viel Zeit genommen hast, meine Daten zu diskutieren.

Großer Dank geht auch an Prof. Dr. Stephan Becker. Lieber Stephan, vielen Dank, dass ich die virologischen Methoden in deiner Arbeitsgruppe durchführen durfte und vielen Dank für die hilfreichen Diskussionen, deine stete Ermutigung, meine Daten auf internationalen Konferenzen zu präsentieren und dein förderliches Feedback.

Vielen Dank an Prof. Dr. Arnold Grünweller und Prof. Dr. Maike Petersen, die sich so kurzfristig bereit erklärt haben, Mitglieder meiner Prüfungskommission zu werden.

Ein Dankeschön an die AG Becker (Nadine, Cornelius, Sandro, Yuki, Nelly, Alexander, Alexandra, Katharina K., Katharina G., Dirk, Astrid, Olga und Verena), die ich immer mit Fragen löchern durfte. Unvergessen bleibt die NSV-Tagung in Verona mit einigen von euch! Insbesondere möchte ich mich bei meinen lieben Laborkollegen Nelly, Anke und Katharina bedanken. Danke, dass für mich immer ein „Miet“-Platz an der Bench oder der Flow frei war. Danke für die tolle Stimmung und gute Musik und die Hilfe beim Zurechtfinden in fremden Laborschubladen und -räumen ;). Nadine, danke dir, dass du mich methodisch eingearbeitet und in das Arbeitsgruppenteam integriert hast und danke für die hilfreichen Diskussionen und Tipps zu Daten und Klonierungen! Danke Dirk und Astrid für das Bereitstellen von Plasmiden oder Zellen.

Ebenfalls bedanken möchte ich mich bei allen Studenten, die mir während ihrer Vertiefungspraktika assistiert haben und insbesondere bei Hannes Huber, der mich im Rahmen seiner Masterarbeit ein Jahr lang tatkräftig unterstützt hat.

Ein ganz besonderer Dank geht natürlich an meine liebe Hartmann-Arbeitsgruppe (Markus, Falk, Laura, Marietta, Clemens, Katrin, Sweetha, Dennis, Sören, Jana W., Amri, Nina, Isabel, Kerstin, Arnold, Dominik, Becki, Katrin, Katja, Nadine, Aileen, Clara, Paul, die Ebola-Crew: Jana D. und Wiebke): Ihr seid die besten! Ohne euch, wären der Laboralltag und die Mittagspausen nur halb so angenehm und lustig gewesen. Katrin, danke dir, dass du mich für mein Praxisjahr in die AG Hartmann geholt, im

experimentellen Arbeiten geschult und für die Wissenschaft begeistert hast. Markus, Marcus und Nadine, danke für euer stets offenes Ohr und eure hilfreichen (u.a. methodischen) Tipps. Arnold und Kerstin, danke für die äußerst wertvollen Anmerkungen, Diskussionen und Unterstützung bei Experimenten. Wiebke, dir möchte ich für das Korrekturlesen, das Auffinden von alten Klonierungsdateien und für die zahlreichen Diskussionen zu Methodenoptimierungen und neuen Daten danken. Danke auch an die allerbeste Laborgang, Dominik und Sweetha. Ich werde die liebevollen Foppereien mit euch vermissen (Sweetha, K-Pop als musikalische Pipettieruntermalung werde ich wohl nicht vermissen :-P).

Nicht zuletzt geht mein Dank an meine Freunde, meine Familie und an Fabian, für eure bedingungslose Unterstützung und Förderung während der gesamten Zeit. Maike, dir danke ich für das Korrekturlesen und dass du beim Lesen des fachfremden Textes tapfer durchgehalten hast! Besonderer Dank gilt dir, Fabian, für deinen Rückhalt, deine unermüdliche Geduld und deine Begleitung in den Tief- und Hochphasen meiner Promotion und im Privaten.